Available online at www.sciencedirect.com
SCIENCE DIRECT®
Molecular Phylogenetics and Evolution 31 (2004) 66–88
MOLECULAR PHYLOGENETICS AND EVOLUTION
www.elsevier.com/locate/ympev

Molecular systematics of the damselfishes (Teleostei: Pomacentridae): Bayesian phylogenetic analyses of mitochondrial and nuclear DNA sequences

Brice Quenouille,[a,b,*] Eldredge Bermingham,[a] and Serge Planes[b]

[a] *Smithsonian Tropical Research Institute, Apdo. 2072, Balboa, Panama*
[b] *Ecole Pratique des Hautes Etudes, UMR CNRS 8046, Université de Perpignan, Perpignan Cédex 66860, France*

Received 17 January 2003; revised 11 July 2003

Abstract

Damselfishes in the family Pomacentridae represent one of the few families of reef fishes found on coral reefs irrespective of location. At a local scale, damselfishes are often the most abundant coral reef fish, and their study has provided much of our current understanding of the ecology of tropical reef animals. The study of phylogenetic relationships among the Pomacentridae has lagged ecological investigation of the group, thus limiting historical perspective on the remarkable species richness of the family. In this study, we used 1989 bp of DNA sequence representing three mitochondrial genes and 1500 bp of the single copy nuclear RAG1 region to infer hypotheses of relationship for the group. Our analysis includes 103 Pomacentridae species in 18 genera, and three of the four named subfamilies: Amphriprioninae, Chrominae, and Pomacentrinae. The Bayesian method of phylogenetic reconstruction was applied to the data, because even with a large number of sequences it is an efficient means of analysis that provides intuitive measures of support for tree topologies and for the parameters of the nucleotide substitution model. Four Pomacentridae clades were identified with high statistical support whether the data were analyzed from a mtDNA, RAG1 or combined perspective, and in all analyses the current subfamilial classification of the Pomacentridae was rejected. At the genus level, *Amphiprion*, *Chromis*, and *Chrysiptera* were also rejected as natural groups. *Abudefduf*, *Amblyglyphidodon*, *Dascyllus*, *Neoglyphidodon*, *Neopomacentrus*, and *Pomacentrus* were each strongly supported as monophyletic genera but the support for monophyly is nonetheless compromised by sample size, except in the case of *Dascyllus* and *Abudefduf* for which we have sampled almost all of the described species.
© 2003 Elsevier Inc. All rights reserved.

Keywords: Pomacentridae; mtDNA; ncDNA; Bayesian methods; Taxonomic sampling

1. Introduction

Damselfishes in the family Pomacentridae number over 350 species and belong to one of the four most speciose families of coral reef fishes (Allen, 1991; Nelson, 1994; Robertson, 1998). They are one of the few reef fish families found on a coral reef regardless of its biogeographic location (Bellwood, 1997), and at local scale damselfishes are often the most abundant fish on a reef (Bellwood, 1996; Myers, 1991). Although recent morphological and molecular analyses have centered renewed attention on the phylogenetic systematics of these fishes (Bernardi and Crane, 1999; Elliott et al., 1999; Fitzpatrick, 1992; Jang-Liaw et al., 2002; McCafferty et al., 2002; Tang, 2001), considerable investigation remains to test the ideas and classification synthesized by Allen (1975, 1991). Furthermore, historical perspective on relationships among damselfishes provides a baseline from which to study the interaction of history and ecology, and the relative roles that evolutionary and ecological processes play in the origin and maintenance of Pomacentridae species diversity (Ricklefs and Schluter, 1993).

Damselfishes were among the first fishes described by Carole Linnaeus (1758) and have subsequently received the attention of other well-known taxonomists (Bleeker, 1877; Cuvier and Valenciennes, 1830). More recently

* Corresponding author. Fax: +33-4-68-50-36-86.
E-mail address: quenovib@naos.si.edu (B. Quenouille).

1055-7903/$ - see front matter © 2003 Elsevier Inc. All rights reserved.
doi:10.1016/S1055-7903(03)00278-1

Allen (1975, 1991) has undertaken an exhaustive effort to describe the taxonomy and geographic distribution of all damselfishes, and to propose a classification for the group. In his last revision (Allen, 1991), 321 species, 28 genera were recognized and classified in the four subfamilies he previously named (Allen, 1975). Since then, a new genera (Allen, 1999a) and 34 new species have been described (Allen, 1992, 1993, 1994, 1999a,b, 2002; Allen and Adrim, 1992, 2000; Allen and Bailey, 2002; Allen and Rajasuriya, 1995; Allen and Randall, 2002; Allen and Smith, 1992; Gasparini et al., 1999; Leão de Moura, 1995; Lessios et al., 1995; Novelli et al., 2000; Randall, 1994, 2001, 2002; Randall and Earle, 1999; Randall and McCosker, 1992). Pomacentridae's taxonomic classification is based entirely on external phenotype, principally body shape, color pattern, and meristic characteristics such as fin ray and gill raker counts. Fitzpatrick (1992) performed the only explicit phylogenetic analysis of the Pomacentridae using morphology and found characters supporting the monophyly of the family, but subfamilial and generic relationships were almost completely unresolved. Overall, the family Pomacentridae appears relatively bereft of phylogenetically informative morphological characters (Emery, 1973; Fitzpatrick, 1992; Gluckmann and Vandewalle, 1998), which is not surprising given that the high species diversity in the family is associated with a relatively small number of general body plans. Molecular systematic investigation of the family is increasing in importance, but most studies to date have focused on congeneric relationships (*Dascyllus*: Bernardi and Crane, 1999; McCafferty et al., 2002; *Amphiprion* + *Premnas*: Elliott et al., 1999; *Abudefduf*: Bermingham et al., submitted; and *Stegastes*: Lessios et al., in prep.). The exceptions are the broader systematic investigations of pomacentrid relationships carried forward by Tang (2001) and Jang-Liaw et al. (2002). Tang (2001) used partial mitochondrial ribosomal (12S, 16S) and tRNA-Phe genes (1471 bp in total) to infer relationships among 23 species representing 14 genera. Jang-Liaw et al. (2002) focused on partial 12S sequences (1058 bp) to present a hypothesis of relationship for 48 pomacentrid species representing 18 genera. Notwithstanding the long history of investigation of Pomacentridae, none of the broad surveys of pomacentrid relationships, whether based on morphology or molecular data, have developed a sufficient number of characters for robust inference of relationships among genera and subfamilies.

Our molecular systematic study of Pomacentridae builds on previous molecular systematic evaluation of the family in three important directions: (1) we have expanded the number of taxa analyzed; (2) a protein-coding nuclear gene has been added to the analysis; and (3) we have re-directed the mitochondrial focus to protein-coding genes. Our taxonomic sample includes 103 pomacentrid species representing 18 of the 29 pomacentrid genera currently recognized (Allen, 1991, 1999a). The 11 missing genera comprise one or two species each and thus represent a minor contribution to species diversity in the family. Six of the missing genera are moderately widespread in the western Pacific and thus might be important to a full understanding of Pomacentridae diversification history, but the other five represent single island endemics at the periphery of damselfish distribution. Our investigation includes almost complete taxonomic sampling of *Dascyllus* and *Abudefduf*, the monotypic *Acanthochromis*, *Hemiglyphidodon*, and *Premnas* genera, and another 13 genera are represented by 20–50% of their species count. The phylogenetic analysis of the Pomacentridae presented here can be combined with all previously published molecule-based analyses of species in the family to provide molecular systematic hypothesis for the family that covers 120 species (34%) and 21 genera (72%).

Our phylogenetic analysis is based on 1992 nucleotides representing the mitochondrial ATP synthase 8 and 6 (ATPase 8 and ATPase 6) and cytochrome *b* (cyt *b*) genes, and a 1500 bp intronless fragment of the nuclear-encoded, single copy recombinant activating gene 1 (RAG1). We utilized mitochondrial protein-coding genes located on opposite sides of the circular mtDNA in order to guard against inadvertent analysis of a mitochondrial pseudogene translocated to the nucleus (see Bermingham et al., 1996; Lopez et al., 1994). The nuclear RAG1 provides a second, unlinked genetic marker of phylogenetic relationship, thus permitting independent inference of both the branching order and time of pomacentrids. In a recent comparison of the phylogenetic utility of three ncDNA gene regions, RAG1 proved to be the most useful in terms of substitution rate and sequencing efficiency (Quenouille and Bermingham, submitted). In total our phylogenetic analysis is based on 3.5 kb of mtDNA and ncDNA sequence data and permits reasonable assessment of problems in phylogenetic resolution that may owe to discordant gene histories or nucleotide saturation. Our phylogenetic reconstruction of Pomacentridae relationships relies predominantly upon Bayesian methods of analysis (Huelsenbeck et al., 2001a; Larget and Simon, 1999). Given the moderately large number of Pomacentridae DNA sequences presented here, Bayesian methods provide a computationally efficient method of analysis, as well as intuitive measures of support for tree topologies and for the parameters of the complex nucleotide substitution models that seem most appropriate to our analysis of damselfish relationships.

Our aims are 2-fold with regard to the systematics of the Pomacentridae. First, we wish to contribute to an improved systematic understanding of relationships within the family and advance knowledge regarding the relative efficacy of different molecular markers of

phylogeny. Second, given the taxonomic holes in the phylogenetic studies of the Pomacentridae to date, we have placed our results in a context that utilizes geographic distribution and prevailing taxonomy to infer the likely effect of missing taxa on the stability of our phylogenetic hypothesis for the family. We anticipate that the composite phylogenetic impression presented here will direct consideration of the best means to advance our understanding of the diversification history of the Pomacentridae. In turn, a similar approach applied to other families of coral reef fishes will provide the historical data that is fundamental to understanding the evolutionary origin and maintenance of the extraordinary diversity of life that inhabits coral reefs.

2. Materials and methods

2.1. Taxon sampling and DNA extraction

One hundred and seventy one individuals were analyzed, representing 103 pomacentrid species in 18 genera and single representatives of each of the other three families included in the suborder Labroidei: Embiotocidae, Labridae, and Cichlidae. For *Abudefduf* and *Dascyllus,* our analysis includes representative individuals and species based on our previous phylogenetic analyses of species relationships in these genera (Bermingham et al., submitted; McCafferty et al., 2002). Species names and collecting locales are reported in Table 1. Specimens were collected by micro-spear, nets or rotenone poisoning. Fish were maintained on ice until we excised a small piece of gill tissue, which was subsequently placed and stored in DMSO/NaCl buffer. DNA voucher specimens representing 31% of the species and 61% of the genera analyzed in this study were preserved in buffered formalin, transferred to 70% ethanol, and deposited in the permanent fish collection at the Smithsonian Tropical Research Institute (STRI: Bermingham et al., 1997). Total genomic DNA was extracted from 0.1 to 0.5 g of gill tissue using a standard phenol/chloroform extraction procedure. The DNA was collected by ethanol precipitation, washed in 70% ethanol (v/v), re-suspended in a 50–150 μl volume of 1/10 TE (1 mM Tris and 0.1 mM EDTA), and stored at −80 °C.

2.2. DNA amplification and sequencing

The mitochondrial ATP synthase 8 and 6 genes were simultaneously amplified with the primers L8331 (5′-AAAGCRTYRGCCTTTTAAGC-3′ (Meyer, 1993)) and H9236 (5′-GTTAGTGGTCAKGGGCTTGGRTC-3′, modified from Meyer (1993)). PCR amplifications were performed on a MJ research PTC-200 thermocycler, in a 25 μl reaction containing: 2.5 μl of 10× buffer (Perkin–Elmer, Norwalk, CT), 200 μM of each dNTP, 0.5 μM of each primer, 2.5 U Amplitaq polymerase (Perkin–Elmer, Norwalk, CT), 2.0 mM $MgCl_2$, and 1–2 μl of total genomic DNA. The thermocycle procedure (35 cycles) was a modified hot-start touchdown PCR, with an initial soak at 94 °C for 3 min, followed by 10 cycles at 94 °C for 45 s, 56 °C for 45 s and 72 °C for 45 s, with a reduction of 0.5 °C in annealing temperature for each cycle (from 56 to 51 °C). This was followed by 25 additional cycles using the same conditions except that the annealing temperature was fixed at 51 °C. The final phase of the procedure was a 5-min elongation period at 72 °C.

The cytochrome *b* gene was amplified with the primers GLUDG-5′ (5′-TGACTTGAARAACCAYCGTTG-3′ Palumbi, 1996) and H16460 (5′-CGAYCTTCGGATTACAAGACCG-3′ (http://nmg.si.edu/bermlab.htm)). We used a modified hot-start touchdown PCR (35 cycles) with an initial soak at 94 °C for 3 min, followed by 20 cycles at 94 °C for 45 s, 58 °C for 45 s, and 72 °C for 1 min 15 s, with a reduction of 0.4 °C in annealing temperature for each cycle (from 58 to 50 °C). This was followed by 15 additional cycles with the annealing temperature fixed at 50 °C. The final cycle finished with a 5-min elongation period at 72 °C.

Amplification of a 1500-bp fragment of the single copy ncDNA RAG1 was performed with the primers RAG1F (5′-AGCTGTAGTCAGTAYCACAARATG-3′, Quenouille and Bermingham, submitted) and RAG9R (5′-GTGTAGAGCCAGTGRTGYTT-3′, Quenouille and Bermingham, submitted). We used a modified hot-start touchdown PCR (35 cycles), with an initial soak at 94 °C for 3 min, followed by 10 cycles at 94 °C for 45 s, 58 °C for 45 s, and 72 °C for 1 min 15 s, with a reduction of 0.5 °C in annealing temperature for each cycle (from 58 to 53 °C). This was followed by 25 additional cycles with the annealing temperature fixed at 53 °C, and the final cycle ended with a 5-min elongation period at 72 °C.

Amplification products were visualized by running 2 μl of the PCR product on a 1% TBE agarose gel (Nusieve GTG, FMC) in the presence of EtBr. When positive, the remaining 23 μl of each reaction were run on a 2% low-melting TAE agarose gel (NuSieve GTG, FMC) in the presence of EtBr. The amplification product was cut from the gel, melted at 70 °C for 5 min, and then digested with Gelase (Epicentre Technologies, Madison WI) for at least 3 h. Two to five microliters of the purified PCR product was used as template in a 10-μl cycle sequencing reaction containing 1 μl of primer, 1–2 μl of dye-terminator reagent (dRhodamine, Applied Biosystems), and H_2O. In all cases the amplification primers listed above were used to sequence the ends of the PCR products and the following internal primers were used to complete the sequences: 8.3 L8524 (5′-AAYCCTGARACTGACCATG-3′, http://nmg.si.edu/bermlab.htm) for the ATP synthase 8 and 6 genes; CB3-3′

Table 1
Species identifications and collecting site locations for the Pomacentridae individuals and outgroup taxa analyzed for this study

Taxa	*n*	Collecting locales	STRI ID	Catalogue nbr	ATPase 8/6	Cytochrome *b*	RAG1
Pomacentridae							
Amphiprioninae							
Amphiprion akallopisos	2	Mozambique	**stri-x-1578,** stri-x-1579		AY208344, AY208345	AY208508	
A. akindynos	2	Australia	stri-x-1591, **stri-x-1592**		AY208346, AY208347	AY208509	AY208628
A. allardi	2	Mozambique	**stri-x-1601,** stri-x-1602		AY208348, AY208349	AY208510	AY208629
A. chrysogaster	2	Mauritius	**stri-x-1610,** stri-x-1611		AY208350, AY208351	AY208511	
A. chrysopterus	2	Guam	**stri-x-1618,** stri-x-1619		AY208352, AY208353	AY208512	
A. clarkii	2	Malaysia, Gulf of Oman	**stri-x-2799,** stri-x-2801		AY208354, AY208355	AY208513	
A. melanopus	2	Australia (GBR), Guam	**stri-x-1662,** stri-x-1671		AY208356, AY208357	AY208514	
A. nigripes	1	Aquarium traders	**stri-x-2809**		AY208358	AY208515	
A. ocellaris	2	Indonesia	**stri-x-2803,** stri-x-2804		AY208359, AY208360	AY208516	AY208631
A. omanensis	2	Gulf of Oman	**stri-x-2882,** stri-x-2883		AY208361, AY208362	AY208517	
A. perideraion	2	Australia (GBR), NC	stri-x-1715, **stri-x-1717**		AY208363, AY208364	AY208518	AY208630
A. sebae	2	Indonesia	**stri-x-2805,** stri-x-2806		AY208365, AY208366	AY208519	
P. biaculeatus	1	Malaysia	**stri-x-2808**		AY208367	AY208520	AY208632
Chrominae							
A. polyacanthus	2	Australia (GBR)	**stri-x-1541,** stri-x-1549		AY208368, AY208369	AY208521	AY208625
Chromis agilis	2	Johnston Atoll, NC	**stri-10709,** stri-x-1765	stri-5562	AY208370, AY208371	AY208522	AY208641
C. amboinensis	2	NC	**stri-10748,** stri-10749	stri-5563	AY208372, AY208373	AY208523	
C. atrilobata	2	Panama	**stri-x-1815,** stri-x-1820		AY208374, AY208375	AY208524	AY208637
C. atripectoralis	2	Australia (GBR), NC	**stri-10686**	stri-5564	AY208376	AY208525	
C. atripes	2	NC	**stri-10736,** stri-10751	stri-5565	AY208377, AY208378	AY208526	AY208633
C. chromis	2	France	**stri-x-2768,** stri-x-2769		AY208379, AY208380	AY208527	AY208640
C. chrysura	2	NC	**stri-10700,** stri-10737	stri-5567	AY208381, AY208382	AY208528	
C. cyanea	2	Jamaica	**stri-9869,** stri-9870	stri-5568	AY208383, AY208384	AY208529	AY208639
C. flavomaculata	2	NC	**stri-10785,** stri-10786	stri-5569	AY208385, AY208386	AY208530	
C. iomelas	2	French Polynesia	**stri-10558,** stri-x-1833	stri-5570	AY208387, AY208388	AY208531	
C. margaritifer	2	NC	**stri-10559,** stri-10560	stri-5571	AY208389, AY208390	AY208532	
C. multilineata	2	Jamaica	stri-9886, **stri-9887**		AY208391, AY208392	AY208533	AY208636
C. nitida	2	Australia	**stri-x-1899,** stri-x-1900		AY208393, AY208394	AY208534	AY208638
C. retrofasciata	2	NC	**stri-10758,** stri-10760	stri-5572	AY208395, AY208396	AY208535	AY208634
C. viridis	2	Japan	stri-9215, **stri-10337**	stri-5573	AY208397, AY208398	AY208536	AY208635
C. weberi	2	Australia (GBR)	**stri-x-2925,** stri-x-2926		AY208399, AY208400	AY208537	AY208642
C. xanthopterigya	2	Gulf of Oman	stri-x-2904, **stri-x-2906**		AY208401, AY208402	AY208538	
Dascyllus aruanus	1	NC	**stri-x-2103**		AF489764	AY208539	AY208649
D. carneus	1	Kenya	**stri-x-2159**		AF489770	AY208540	AY208652
D. flavicaudus	1	French Polynesia	**stri-x-2188**		AF489775	AY208541	AY208650
D. marginatus	1	Red Sea	**stri-x-2218**		AF489782	AY208542	AY208651
D. melanurus	1	Phillipines	**stri-x-2242**		AF489788	AY208543	AY208653
D. reticulatus	1	Australia (GBR)	**stri-x-2250**		AF489739	AY208544	AY208654
D. trimaculatus	1	Maldives	**stri-x-2345**		AF489820	AY208545	AY208655
Pomacentrinae							
Abudefduf abdominalis	1	Hawaii	**stri-x-898**		AY208403	AY208546	
A. bengalensis	1	Indonesia	**stri-x-928**		AY208404	AY208547	
A. concolor	1	Panama	**stri-x-995**		AY208405	AY208548	

Table 1 (*continued*)

Taxa	*n*	Collecting locales	STRI ID	Catalogue nbr	ATPase 8/6	Cytochrome *b*	RAG1
A. declivifrons	1	Mexico	**stri-x-1043**		AY208406	AY208549	
A. lorenzi	1	Indonesia	**stri-x-1049**		AY208407	AY208550	
A. margariteus	1	Reunion	**stri-x-6123**		AY208408	AY208551	AY208618
A. notatus	1	Indonesia	**stri-x-6127**		AY208409	AY208552	
A. saxatilis	1	Panama	**stri-x-1094**		AY208410	AY208553	AY208624
A. septemfasciatus	1	French Polynesia	**stri-x-1161**		AY208411	AY208554	AY208619
A. sexfasciatus	1	Australia (GBR)	**stri-x-1193**		AY208412	AY208555	AY208620
A. sordidus	1	Ascension Isl.	**stri-x-1251**		AY208413	AY208556	AY208621
A. sparoides	1	Reunion	**stri-x-6101**		AY208415	AY208558	
A. taurus	1	Panama	**stri-x-1310**		AY208416	AY208559	AY208622
A. troschelii	1	Panama	**stri-x-1438**		AY208417	AY208560	AY208623
A. vaigiensis	1	Christmas Island	**stri-x-6064**		AY208418	AY208561	
A. whitleyi	1	Australia (GBR)	**stri-x-1535**		AY208419	AY208562	
Amblyglyphidodon aureus	2	Australia (GBR)	stri-x-2919, **stri-x-2920**		AY208420, AY208421	AY208563	AY208627
A. curacao	2	Australia (GBR), NC	stri-10531, **stri-x-1563**	stri-5561	AY208423, AY208422	AY208564	
A. leucogaster	2	NC	stri-10544, **stri-10819**	stri-5575	AY208424, AY208425	AY208565	AY208626
C. annulata	2	Gulf of Oman	**stri-x-2867,** stri-x-2868		AY208426, AY208427	AY208566	AY208645
C. cyanea	2	Japan	**stri-9204,** stri-9205	stri-5577	AY208428, AY208429	AY208567	AY208643
C. galba	2	Indonesia	stri-x-2813, **stri-x-2814**		AY208430, AY208431	AY208568	
C. glauca	2	French Polynesia	**stri-x-1957,** stri-x-1958		AY208432, AY208433	AY208569	AY208647
C. hemicyanea	1	Indonesia	**stri-x-2815**		AY208434	AY208570	AY208644
C. brownrigii	2	French Polynesia	**stri-x-1971,** stri-x-1972		AY208435, AY208436	AY208571	AY208648
C. rex	2	NC	**stri-10456,** stri-10457	stri-5578	AY208437, AY208438	AY208572	
C. rollandi	2	NC	**stri-10809,** stri-10810	stri-5579	AY208439, AY208440	AY208573	AY208646
C. talboti	2	Australia (GBR)	**stri-x-2931,** stri-x-2932		AY208441, AY208442	AY208574	
C. taupou	2	NC	**stri-10251,** stri-10346	stri-5580	AY208443, AY208444	AY208575	
Dischistodus melanotus	1	Australia (GBR)	**stri-x-2934**		AY208445	AY208576	AY208656
Hemiglyphidodon plagiometopon	2	NC	**stri-10547,** stri-10548	stri-5587	AY208446, AY208447	AY208577	
Microspathodon chrysurus	2	Jamaica	stri-9873, **stri-9916**		AY208448, AY208449	AY208578	AY208657
Neoglyphidodon melas	1	Australia (GBR)	**stri-x-2454**		AY208450	AY208579	
N. nigroris	2	Australia (GBR)	**stri-x-2849,** stri-x-2850		AY208451, AY208452	AY208580	AY208658
N. oxyodon	1	Aquarium traders	**stri-x-2821**		AY208453	AY208581	AY208659
N. polyacanthus	1	NC	**stri-10290**	stri-5582	AY208454	AY208582	
Neopomacentrus cyanomos	2	Australia (North)	**stri-x-2844**		AY208455	AY208583	

N. filamentosus	2	Australia	stri-x-2824, **stri-x-2825**		AY208456, AY208457	AY208584	AY208661
N. miryae	2	Gulf of Oman	**stri-x-2885,** stri-x-2886		AY208458, AY208459	AY208585	
N. nemurus	2	NC	**stri-10600,** stri-10601	stri-5583	AY208460, AY208461	AY208586	
N. sindensis	2	Gulf of Oman	**stri-x-2895,** stri-x-2896		AY208462, AY208463	AY208587	AY208660
Parma oligolepis	1	NC	**stri-10320**	stri-5584	AY208464	AY208588	AY208662
Plectroglyphidodon dickii	2	NC	stri-10564, **stri-10744**	stri-5585	AY208465, AY208466	AY208589	AY208663
P. leucozonus	2	Gulf of Oman	**stri-x-2876,** stri-x-2877		AY208467, AY208468	AY208590	
Pomacentrus adelus	2	NC	**stri-10269,** stri-10270	stri-5588	AY208469, AY208470	AY208591	AY208664
P. australis	2	Australia	stri-x-2477, **stri-x-2478**		AY208471, AY208472	AY208592	AY208668
P. bankanensis	1	NC	**stri-10651**	stri-5589	AY208473	AY208593	AY208665
P. brachialis	2	Australia	**stri-x-2479,** stri-x-2480		AY208474, AY208475	AY208594	
P. chrysurus	1	Australia	**stri-x-2482**		AY208476	AY208595	
P. coelestis	1	Indonesia	**stri-10725,** stri-x-2823	stri-5590	AY208478, AY208477	AY208596	
P. grammorhynchus	2	Australia (GBR)	**stri-x-2946,** stri-x-2948		AY208479, AY208480	AY208597	
P. lepidogenys	2	NC	**stri-10652,** stri-10655	stri-5591	AY208481, AY208482	AY208598	
P. leptus	2	Gulf of Oman	**stri-x-2879,** stri-x-2880		AY208483, AY208484	AY208599	
P. milleri	1	Australia	**stri-x-2847**		AY208485	AY208600	
P. moluccensis	2	Australia	**stri-x-2483,** stri-x-2484		AY208486, AY208487	AY208601	AY208669
P. nagasakiensis	2	Japan, NC	stri-x-9229, **stri-10165**	stri-5592	AY208489, AY208488	AY208602	
P. nigromanus	2	Australia	**stri-x-2853,** stri-x-2859		AY208490, AY208491	AY208603	AY208671
P. pavo	2	French Polynesia, NC	**stri-10657,** stri-x-2487	stri-5593	AY208492, AY208493	AY208604	AY208666
P. philippinus	2	NC	**stri-10717,** stri-10718	stri-5594	AY208494, AY208495	AY208605	AY208667
P. reidi	2	Australia (GBR)	**stri-x-2949,** stri-x-2951		AY208496, AY208497	AY208606	AY208672
P. smithi	2	NC	**stri-10326,** stri-10327	stri-5595	AY208498, AY208499	AY208607	AY208670
P. trilineatus	1	Gulf of Oman	**stri-x-2891**		AY208500	AY208608	
P. vaiuli	2	NC	**stri-10693,** stri-10719	stri-5596	AY208501, AY208502	AY208609	
Pomachromis fuscidorsalis	1	French Polynesia	**stri-x-2503**		AY208503	AY208610	
Stegastes planifrons	1	Panama	**stri-9151**	stri-5598	AY208504	AY208611	AY208673
Embiotocidae							
Embiotoca lateralis	1	Gulf of Mexico	**stri-x-2788**		AY208505	AY208612	AY208615
Labridae							
Halichoeres melanurus	1	NC	**stri-10169**	stri-5599	AY208506	AY208613	AY208617
Cichlidae							
Amphilophus rhytisma	1	Panama	**stri-213**	stri-04240	AY208507	AY208614	AY208616

STRI IDs reported in bold correspond to individuals for which the ATPase8/6, cytochrome b genes and RAG1 (when sequenced) have been analyzed. GBR: Great Barrier Reef, NC: New Caledonia.

(5′-GGCAAATAGGAARTATCATTC-3′, Palumbi, 1996) and CytB425F (5′-TGAGGCCAAATRTMTTYT GAGG-3′, Quenouille and Bermingham, submitted) for the cytochrome *b* gene; RAG3F (5′-GGGTGATGTCA GYGAGAAGCA-3′, Quenouille and Bermingham, submitted) RAG5R (5′-TRGAGTCACACAGACTGC AGA-3′, Quenouille and Bermingham, submitted), and RAG8R (5′-CGCCACACAGGYTTCATCT-3′, Quenouille and Bermingham, submitted) for the RAG1 gene. Cycle sequencing reactions were performed on a MJ research PTC-200 thermocycler following the procedure recommended by Applied Biosystems. Reactions were cleaned of excess nucleotides using Sephadex columns (Princeton Separations), dried, re-suspended in 1.3 μl of formamide-EDTA loading dye, loaded onto 5% Long Ranger (FMC) gels, and then analyzed using an Applied Biosystem 377 automatic sequencer.

2.3. Phylogenetic analyses

Chromatograms were edited using Sequencer software (Gene Codes), trimmed of blank ends and primer sequences, and assembled as contiguous fragments (contigs) representing each gene and individual specimen. Contigs representing different individuals and taxa were aligned by eye. The final set of aligned contigs was output as a NEXUS file. MacClade 4.0 (Maddison and Maddison, 2000) and PAUP* 4.0b10 (Swofford, 2002) were used to describe the data. Preliminary phylogenetic analysis demonstrated that multiple individuals representing the same species had virtually identical ATPase 8 and ATPase 6 sequences and formed a monophyletic and distinct lineages as expected given proper species identification. We subsequently sequenced the cytochrome *b* gene for one individual per species. For the 106 individuals sequenced for the mitochondrial ATPase 8, ATPase 6 and cytochrome *b* genes, a partition homogeneity test (Farris et al., 1995) indicated no heterogeneity in the phylogenetic signal among gene partitions (ATPase 8:ATPase 6, $P = 0.72$; ATPase 8:Cyt *b*, $P = 0.34$; and ATPase 6:Cyt *b*, $P = 0.94$). We thus combined the ATPase 8, ATPase 6, and cytochrome *b* sequences in a single data set, which we named the 106mtDNA data set, for subsequent analysis.

Phylogenetic analyses of 106mtDNA provided preliminary insight into relationships among pomacentrid species, particularly those that were closely related, and permitted us to focus attention on a subset of 55 species that we subsequently sequenced for RAG1. We named this nuclear gene partition 55RAG1. Finally we combined the mitochondrial and nuclear data (55combined), given that a partition homogeneity test indicated that the phylogenetic histories of the mtDNA and RAG data were not significantly different ($P = 0.41$).

Bayesian phylogenetic analyses were performed with MrBayes 2.1 (Huelsenbeck and Ronquist, 2001) following the analytical recommendations of Huelsenbeck et al. (2001b). Sequences representing the Embiotocidae, Labridae, and Cichlidae were positioned at the top of the Nexus files. For all three data sets, Modeltest (Posada and Crandall, 1998) identified complex substitution models, the simplest being TrN93 model (Tamura and Nei, 1993) for the 55RAG1 data set. In any event, MrBayes 2.1 only permits a choice of the HKY85 (Hasegawa et al., 1985) or the GTR (Yang et al., 1994) models. The GTR model was adopted for all data sets as it includes all the parameters of nucleotide substitution found in any of the simpler models identified by Modeltest. We used the MrBayes default settings to establish the initial heating values for four Markov chains, and default settings were also used to initially parameterize the GTR + Γ + I model. The four differentially heated Markov chains were initiated from random trees, run simultaneously, and were sampled every 100 cycles.

Preliminary runs were performed to monitor the fluctuating value of the likelihoods of the Bayesian trees, and stationarity was consistently observed before 50,000 generations for the 55RAG1 and 55combined data sets, and before 200,000 generations for the 106mtDNA data set. The Markov chain analyses used to infer Pomacentridae phylogenies were run for 3×10^6 cycles for the data sets based on 55 species and 5×10^6 cycles for the 106mtDNA data set. All sampled trees preceding stationarity were discarded ("burnin" = 500 and 2000, respectively), and the remaining tree samples were used to generate a 50% majority rule consensus tree. The posterior probability of each clade is provided by the percentage of trees identifying the clade and these are true probabilities given the assumptions of the GTR + Γ + I model (Huelsenbeck and Ronquist, 2001). Thus probabilities of 95% or greater were considered significant. Mean, variance, and 95% credibility intervals (CI) for the nucleotide substitution, Gamma, and invariant site parameters values were calculated from the trees sampled after the Markov chain analysis reached stationarity.

We also used Neighbor-Joining (NJ) and parsimony analyses to permit comparative observations regarding support, or lack thereof, for the clades identified in the Bayesian analysis. The mean parameter values obtained from the Bayesian analysis were used as the basis for a NJ analysis of pomacentrid species based on the GTR + Γ + I model; support for the NJ trees was determined using 1000 bootstrap sequence replicates. Maximum Parsimony (MP) analyses were based on heuristic searches with TBR branch swapping. Each substitution possibility was weighted as the inverse of the parameter values calculated from the Bayesian analysis. Support for the MP trees was determined using 1000 bootstrap sequence replicates. The Shimodeira and Hasegawa (1999) test (SH test) was used to estimate the significance of topological differences between the

different data sets and the different analytical methods. We applied the SH test implemented in PAUP*, using RELL optimization and 10,000 bootstrap replicates. Multiple hits in mtDNA sequences were assessed by plotting uncorrected pairwise distances of the mitochondrial data against uncorrected pairwise distances of the nuclear data.

3. Results

3.1. Sequence data patterns

All sequences reported in this paper have been submitted to GenBank and the accession numbers are reported in Table 1. Across all individuals, the ATPase 8 and ATPase 6 genes demonstrated the typical vertebrate pattern with a 10-bp overlap and a combined length of 842 bp. ATPase 8 was 168 bp and ATPase 6 was 684 bp. Cytochrome *b* sequences were 1140 bp long except for the Cichlidae *Amphilophus rhytisma*, which was three bases shorter (1137 bp). The last cichlid cytochrome *b* codon was homologous to the penultimate codon in all other sequences, and we excluded from analysis the final three cytochrome *b* nucleotides for all samples except *A. rhytisma*. The combined mtDNA data set numbered 1989 nucleotides.

The RAG1 primers used in this study generated sequences of 1500 bp for all individuals; this fragment is homologous to the second half of the 3222 bp RAG 1 *Oncorhyncus mykiss* (Rainbow trout) sequence deposited in Genbank (U15663, positions 1549–3049). The presence of clear double peaks at some chromatogram positions, including redundant information on complementary strands, suggested that some nucleotide sites in the RAG1 gene were heterozygous. Double peaks were coded according to the IUB code corresponding to the appropriate two-base ambiguity (no three or four base ambiguities were observed). Across the 1500 bp of RAG1 sequence we scored 95 ambiguities, representing 43 species and a range of 1–5 ambiguities per species. We refer the reader to Quenouille and Bermingham (submitted) for additional information regarding the molecular characterization of RAG1.

Table 2 provides a data summary for the three pomacentrid data sets (106mtDNA, 55RAG11, 55combined) including variable sites, parsimony informative sites, the ratio of changes at first, second and third codon positions, and the *P* values for tests of the homogeneity of base frequencies. Briefly, on a per site basis for the 55 taxa analyzed for both nuclear and mitochondrial sequences, we observed more than twice the number of parsimony informative characters in the mtDNA sequence data as compared RAG1. In terms of codon position both the mitochondrial and nuclear genes carried more variation at third positions, followed by first positions and then second positions, as expected for protein-coding genes.

3.2. Phylogenetic analyses

Preliminary phylogenetic analysis and comparison of mtDNA and RAG1 sequences helped refine our RAG1 sequencing strategy, leading to the subset of 55 taxa

Table 2
Summary statistics resulting from the phylogenetic analyses of 103 Pomacentridae species and three outgroup families

	Mitochondrial (106 sequences)	RAG1 (55 sequences)	mtDNA + RAG1 (55 sequences)
Number of sites	1989	1500	3489
Number of variable sites	1127 (56.6%)	526 (35%)	1629 (46.6%)
Parsimony informative	990 (49.8%)	340 (22.6%)	1297 (37.1%)
variation across codon position (1/2/3 ratio)	2/1/4.15	1.67/1/5.95	1.9/1/4.9
Test for homogeneity of base frequencies	$P = 0.93$ (df = 318)	$P = 1$ (df = 162)	$P = 0.99$ (df = 162)
Π_A	0.311 (0.297–0.323)	0.234 (0.215–0.254)	0.310 (0.299–0.322)
Π_C	0.365 (0.354–0.378)	0.258 (0.239–0.278)	0.341 (0.330–0.352)
Π_G	0.083 (0.079–0.088)	0.288 (0.268–0.309)	0.127 (0.119–0.136)
Π_T	0.240 (0.231–0.248)	0.218 (0.200–0.237)	0.220 (0.212–0.229)
R_{CT}	4.929 (4.163–5.786)	7.589 (5.696–10.212)	6.492 (5.285–7.885)
R_{CG}	0.737 (0.589–0.915)	0.990 (0.696–1.407)	0.643 (0.516–0.814)
R_{AT}	0.534 (0.436–0.656)	1.144 (0.770–1.660)	0.730 (0.563–0.917)
R_{AG}	8.991 (7.681–10.535)	4.986 (3.706–6.698)	5.475 (4.642–6.445)
R_{AC}	0.381 (0.310–0.457)	1.706 (1.188–2.410)	0.507 (0.405–0.636)
Γ	0.665 (0.630–0.704)	0.866 (0.590–1.204)	0.559 (0.518–0.602)
Pinv	0.406 (0.386–0.429)	0.474 (0.390–0.541)	0.463 (0.441–0.484)
Number of trees sampled	48,000	29,500	29,500
Number of distinct topologies	30,873	17,046	964

The middle rows present the mean parameter estimates and 95% confidence intervals determined from the "posterior" distribution of parameter values following Bayesian analysis. The final two rows identify the total number of trees, and the number of distinct tree topologies, sampled from the Bayesian analysis of Pomacentridae relationships.

sequenced for RAG1 in addition to their mitochondrial ATPase and cytochrome *b* genes. As can be determined from Fig. 1, RAG1 has a substitution rate that is approximately 12 times slower than the mitochondrial genes. Thus closely related species were typically not separated by sufficient RAG1 sequence divergence to warrant nuclear gene analysis of the entire set of 106 species. For example, the 15 species pairs in our study separated by 10% or less mtDNA sequence divergence have a maximum level of 0.95% RAG1 divergence. The range for these 15 pairs is 4.01% (80 substitutions) to 10% (190 substitutions) mtDNA divergence in comparison to 0.13% (2 substitutions) to 0.95% (14 substitutions) RAG1 divergence. Below, in the context of specific phylogenetic results, we provide additional detail supporting our sub-sampling strategy for RAG1, but as a general rule, RAG1 was collected for representative lineages of each clade mtDNA sequences identified with strong statistical support.

The mean value and the 95% CI of the GTR + Γ + I model parameters estimated from Bayesian phylogenetic analyses are reported in Table 2 for the 106mtDNA, 55RAG11, and 55combined data sets that served as the basis for our phylogenetic analysis of the Pomacentridae. Across data sets, the 95% CIs are relatively narrow for most parameters excepting the R_{AG} and R_{CT} transition rates, and the α shape parameter in the particular example of the 55RAG1 data set. As expected, base frequencies are different between the mtDNA and ncDNA sequences, with mtDNA sequences evidencing the genome's typical anti-guanine bias ($\Pi_G = 0.087$). It is noteworthy that the transition and transversion parameters estimates made from Bayesian posterior probabilities are consistent with the nucleotide substitution model predicted by Modeltest. For example, the hLRT criteria of Modeltest suggested a TrN93 + Γ + I substitution model for RAG1, and in Table 2 we see that R_{AC}, R_{AT}, R_{CG}, and R_{GT} (in MrBayes R_{GT} is fixed at 1 a priori) have overlapping CIs. These four transversion rate classes are assumed equal in the TrN93 model.

The number of distinct tree topologies sampled after achieving stationarity in the Bayesian analyses is reported in the final rows of Table 2. For each data set the number of distinct trees is large, indicating considerable uncertainty in the phylogenetic signal of the mitochondrial and RAG1 genes whether considered alone or in combination. For the 106mtDNA and the 55RAG1 analyses, nearly every other sampled tree is a different topology. In the sample of trees obtained from the combined data analysis, the number of distinct topologies is much lower (964) indicating a substantial improvement in the phylogenetic resolution of Pomacentridae relationships. Still, further examination of the tree sample based on the 55combined data revealed that the topology with the highest posterior probability represents only 6.4% (1890 trees) of the total sample of 29,500 trees. Furthermore, 302 different topologies (each a distinct hypothesis of Pomacentridae relationship) are required to reach a 95% CI for the group.

Consistent with the Bayesian results, bootstrap consensus trees obtained from NJ and MP analyses revealed a large number of polytomies, and levels of bootstrap support that often fell under 50%. Overall, both marker-specific and combined phylogenetic analyses of 1989 bp of mtDNA sequences and 1500 bp of ncDNA sequences

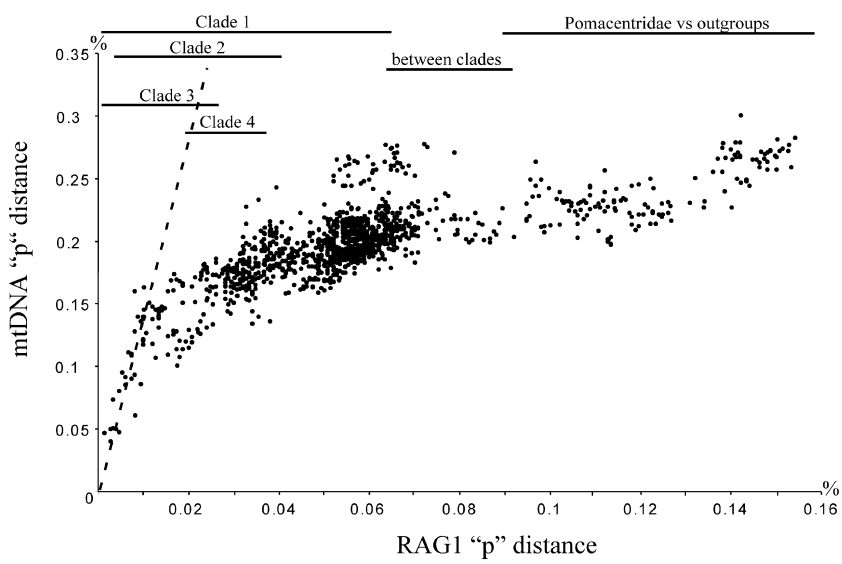

Fig. 1. Pairwise uncorrected ("p") mitochondrial distances plotted against the corresponding uncorrected ("p") nuclear genetic distances. Pairwise comparisons are based on 52 Pomacentridae species and a single species representing each of the three outgroup families: Cichlidae, Labridae, and Embiotocidae. Uncorrected mitochondrial distances were based on 1989 nucleotides of the ATPase 8/6 and cytochrome *b* genes. Uncorrected nuclear distances were based on 1500 nucleotides representing the single copy RAG1 gene.

failed to fully resolve the phylogenetic relationships among Pomacentridae taxa, regardless of analytical method. Nonetheless, as demonstrated by Figs. 2–4, all three data sets produced trees for which many branches received high statistical support using Bayesian, NJ, and MP methods.

There are two overarching phylogenetic results that are consistent across all trees and analyses. First, all treatments of the data suggest that the 103 pomacentrid species analyzed here represent a monophyletic assemblage relative to the other three families typically included with the Pomacentridae in the Labroidei: Cichlidae, Labridae, and Embiotocidae (Figs. 2–4). The statistical support for the monophyletic status of the Pomacentridae is very high (Pr = 1.0 and BP ⩾ 92). Second, we are able to identify four pomacentrid groups, designated Clades 1–4, which have strong statistical support (Pr = 1.0 and BP ⩾ 92), and encompass the same genera and species for all data sets and all analyses (Figs. 2–4). The molecular systematic results presented here fail to support the monophyly of two Pomacentridae subfamilies included in this analysis: Chrominae and Pomacentrinae. A tree that constrained the monophyly of these subfamilies was significantly worse than the consensus trees in a SH test ($P = 0.000$ for all data sets). Additional clades, discussed below, are also consistently identified across all trees, but the four named clades represent the minimum number of groups that associate all analyzed pomacentrid species with a high degree of confidence. Clade 1 includes species representing the following genera: *Acanthochromis*, *Amblyglyphidodon*, *Amphiprion*, *Chrysiptera*, *Dischistodus*, *Hemiglyphidodon*, *Neoglyphidodon*, *Neopomacentrus*, *Pomacentrus*, *Pomachromis*, and *Premnas* (Pr = 1.0, BP = 100). Clade 2 unites all *Chromis* and *Dascyllus* species (Pr = 1.0, BP = 100), and clade 3 comprises only species in the genus *Abudefduf* (Pr = 1.0, BP ⩾ 92). Clade 4 unites species representing the following genera: *Microspathodon*, *Parma*, *Plectroglyphidodon*, and *Stegastes* (Pr = 1.0, BP ⩾ 95).

The Pomacentridae trees pictured in Figs. 2–4 also provide strong support for additional phylogenetic structure within Clades 1–4. Here, we center our focus on those results for which our species sampling within genera is good to moderate. As noted above, Pomacentridae Clade 3 is comprised entirely of *Abudefduf* species, and our taxon sampling includes 16 of the 20 described species in this genus. Within *Abudefduf* there is significant statistical support for at least three natural groups. The eastern Pacific and Caribbean night sergeant majors (*Abudefduf declivifrons*, *Abudefduf concolor*, and *Abudefduf taurus*) are the sister group to all other *Abudefduf* (Pr = 1.0, BP ⩾ 95). In turn, there is strong statistical support for the reciprocal monophyly of the clade comprised of *Abudefduf sordidus*, *Abudefduf septemfasciatus*, and *Abudefduf notatus* and its sister group of all remaining *Abudefduf* (Pr = 1.0, BP ⩾ 100). Lastly, there is strong mitochondrial support for a group of closely related species that includes *Abudefduf sexfasciatus* and the geminate pair *Abudefduf saxatilis* and *Abudefduf troschelii* (Pr = 1.0, BP = 100), but RAG1 provides lower bootstrap support for the monophyly of this lineage (Pr = 1.0, BP = 79). The difference in support values reflects the very small number of phylogenetically informative RAG1 substitutions ($N = 2$) that support the monophyly of *A. sexfasciatus*, *A. saxatilis*, and *A. troschelii*, in contrast to 38 mtDNA synapomorphies that identify this clade.

In Clade 2, the monophyly of *Dascyllus* is strongly supported (Pr = 1.0, BP = 100). Our sample includes nine of the 10 described *Dascyllus* species, and the only species that we failed to sample, *Dascyllus auripinnis*, is considered very closely related to *Dascyllus trimaculatus*. Thus, the monophyly of *Dascyllus* is a very robust result from a mitochondrial perspective that is unlikely to be modified by additional sampling or analysis. Although we do not present the results here, it is worth noting that we sequenced the RAG1 1500 bp gene fragment for all *Dascyllus* species except *D. auripinnis*, and subsequent phylogenetic analysis only resolved the sister group relationship of *Dascyllus melanurus* and *Dascyllus aruanus* to all other *Dascyllus* species. Finer resolution of relationships among *Dascyllus* species was not possible with the 1500 bp RAG1 gene fragment, indicating that the speciation rate was fast relative to the RAG1 substitution rate for this group, a result that parallels that presented above for closely related *Abudefduf* species. RAG1 sequences representing the additional *Dascyllus* species not presented in Fig. 3 have been deposited in GenBank (Table 1).

Chromis, the other genus falling in Clade 2, does not appear to be a natural group. Although our sample of *Chromis* includes only 20% of the described species, the strong support for the clade we designate *Chromis* II (Pr = 1.0, BP = 100), and its sister group relationship to *Dascyllus* (Pr = 1.0, BP ⩾ 92), permit strong inference that increased taxon sampling will not render *Chromis* monophyletic. Furthermore, a tree constrained by the monophyly of *Chromis* was significantly worse than the trees shown in Figs. 2–4, except when the SH test was based on RAG1 data alone (106mtDNA: $P = 0.0007$, 55RAG: $P = 0.031$, 55combined: $P = 0.1207$). The level of support for relationships among the species that comprise *Chromis* I is generally weak at the clade's base, and thus we anticipate that greatly increased species and nucleotide sampling will be needed to interpret the full diversification history of the *Chromis* + *Dascyllus* clade.

Pomacentridae Clade 1 contains roughly 50% of the generic level diversity sampled in our study. Fifty-eight species were analyzed with mtDNA sequences, representing 34% of all species (173) assigned to genera that fall in Clade 1. Consistent with the results reported for

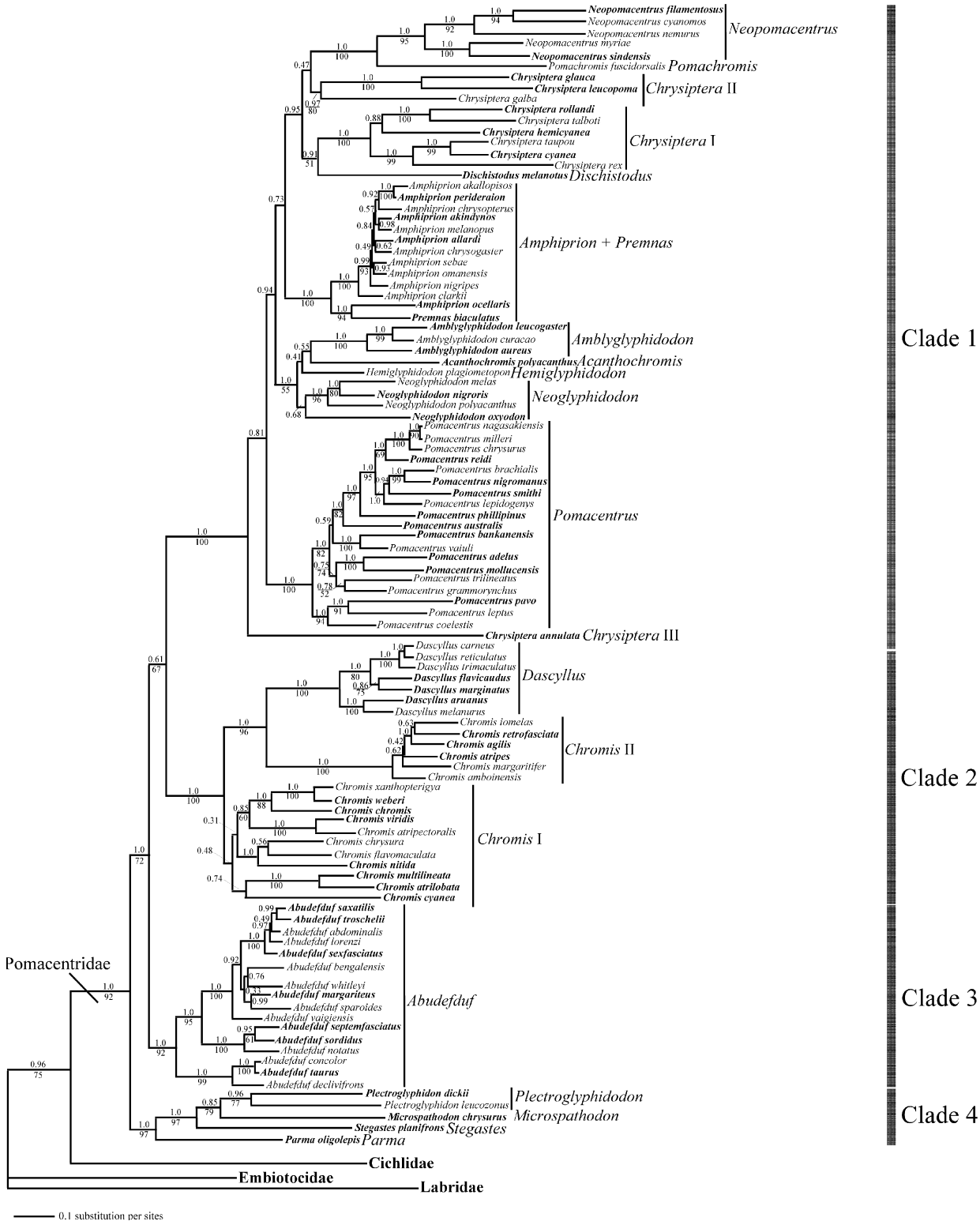

Fig. 2. The 50% majority-rule consensus tree from the Bayesian analysis of the combined mitochondrial ATPase8, ATPase6 and cytochrome *b* genes (total of 1989 bp) representing 103 Pomacentridae and single representatives of the Cichlidae, Labridae, and Embiotocidae. The phylogeny reported corresponds to the consensus topology of 48,000 trees sampled from Bayesian analysis assuming a GTR + Γ + I model (see text for details). Numbers above the branches correspond to posterior probabilities estimated using the Bayesian approach, and numbers below branches refer to the bootstrap support calculated from NJ analysis of 1000 sequence replicates assuming model parameters values estimated from the Bayesian analysis. The phylogenies obtained from NJ and MP analyses were highly congruent with the Bayesian topology. Bootstrap estimates obtained from MP analysis were virtually identical to NJ bootstrap estimates and are not reported. Numbered clades are referred to in the text. The 55 species in bold were selected for additional phylogenetic analysis based on the nuclear RAG1 gene (see text for details).

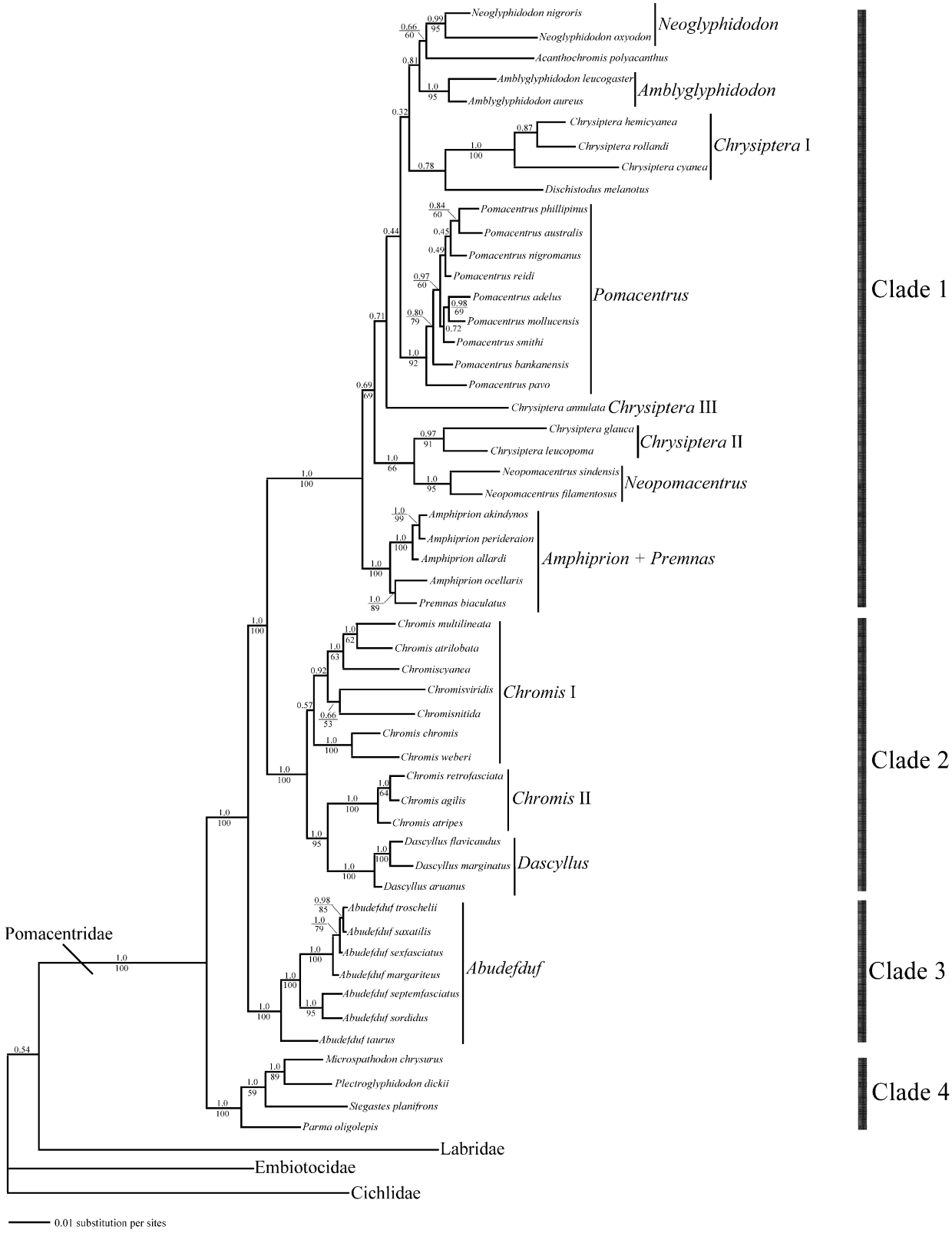

Fig. 3. The 50% majority-rule consensus tree from the Bayesian analysis of 1500 bp of the single copy nuclear gene RAG1 representing 52 Pomacentridae and single representatives of the Cichlidae, Labridae, and Embiotocidae. The 55 species were selected based on their placement in the 106 species mtDNA phylogeny reported in Fig. 2 and preliminary phylogenetic analyses of RAG1 gene (see text for details). The phylogeny corresponds to the consensus topology of 29,500 trees sampled from a Bayesian analysis assuming a GTR + Γ + I model. The phylogenies obtained from NJ and MP analyses were highly congruent with the Bayesian topology and support values above and below branches were calculated as described for Fig. 2.

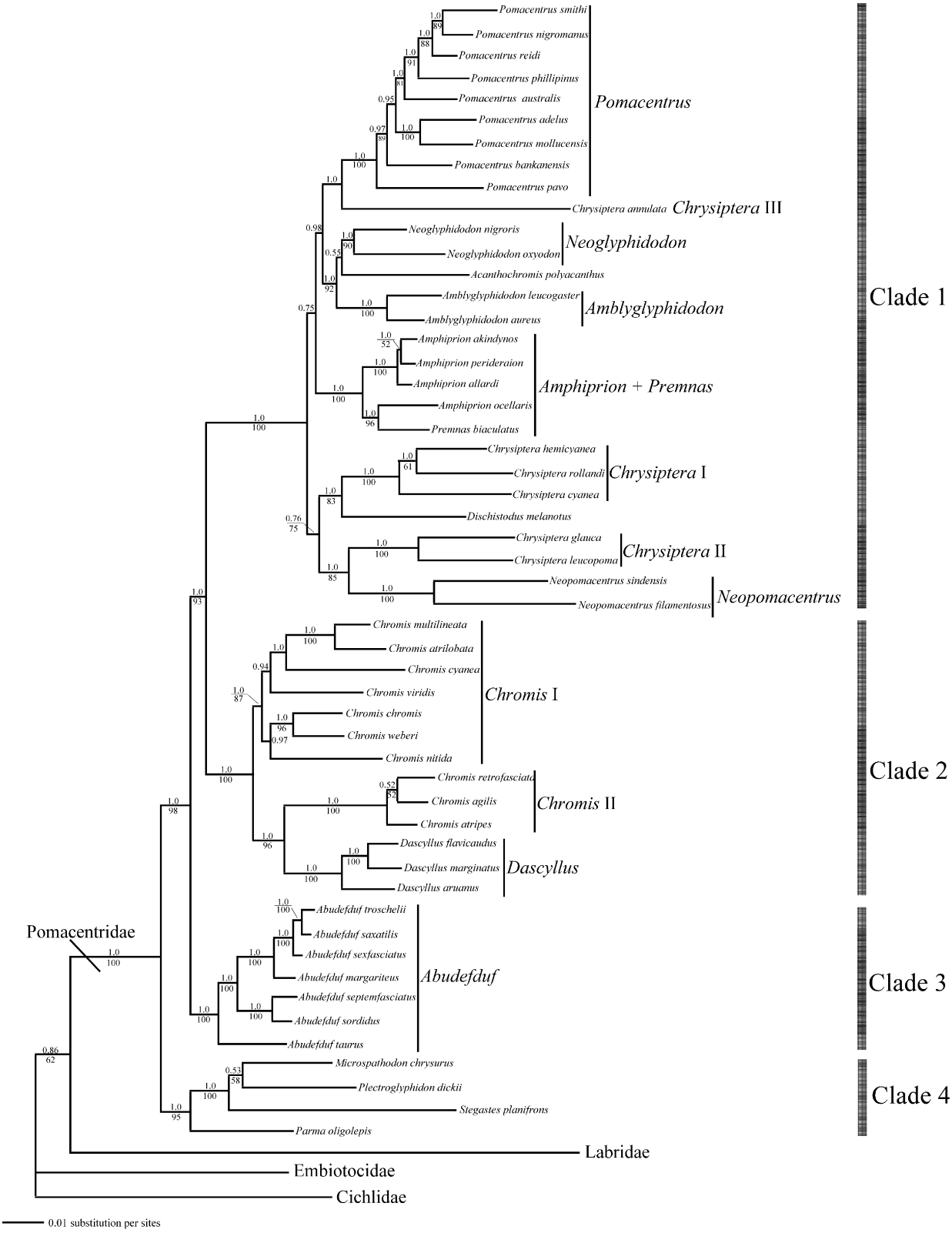

Fig. 4. The 50% majority-rule consensus tree from the Bayesian combined analysis of mitochondrial DNA ATPase 8, ATPase 6 and cytochrome *b* genes (total of 1989 bp) and 1500 bp of the single copy nuclear gene RAG1 representing 52 Pomacentridae and single representatives of the Cichlidae, Labridae, and Embiotocidae. The 55 species were selected based on their placement in the 106 species mtDNA phylogeny reported in Fig. 2 and preliminary phylogenetic analyses of RAG1 gene (see text for details). The phylogeny corresponds to the consensus topology of 29,500 trees sampled from a Bayesian analysis assuming a GTR + Γ + I model. The phylogenies obtained from NJ and MP analyses were highly congruent with the Bayesian topology and support values above and below branches were calculated as described for Fig. 2.

Clade 3 (*Abudefduf*) and Clade 2 (*Dascyllus*), 1500 bp of RAG1 sequence was not sufficiently variable to generate a strong phylogenetic signal of relationships among congeneric species. For example, across the 19 *Pomacentrus* species analyzed, 666 mtDNA positions were variable, of which 417 characters were parsimony informative. Nine of the most divergent species were sequenced for RAG1, which yielded 61 substitutions of which 15 were parsimony informative.

On the other hand, Bayesian, NJ, and MP analyses strongly support most, but not all, genera in Pomacentridae Clade 1 as natural groups. For example, there is clear statistical support for the monophyly of three species groups: *Pomacentrus* (Pr = 1.0, BP ⩾ 92), *Neopomacentrus* (Pr = 1.0, BP ⩾ 95), and *Amblyglyphidodon* (Pr = 1.0, BP ⩾ 95). Our analysis includes roughly 30% of the *Pomacentrus* species, and approximately 36% and 30% of the described species representing *Neopomacentrus* and *Amblyglyphidodon*, respectively, thus providing preliminary molecular systematic support for the traditional view (Allen, 1991) that these three genera represent natural groups. *Neoglyphidodon* also has moderate species representation (50%), and the monophyly of the genus has significant statistical support based on the 55RAG1 and 55combined data sets (Pr ⩾ 0.99, BP ⩾ 95), but not when based on the mtDNA data alone (106mtDNA: Pr = 0.68, BP<50). Lastly within Pomacentridae Clade 1, there was extremely high statistical support for a clade comprising *Premnas biaculeatus* and all *Amphiprion* species (Pr = 1.0, BP = 100). However the monophyly of *Amphiprion* is called into question owing to the sister status of the monotypic *P. biaculeatus* and *Amphiprion ocellaris*, a relationship with strong support across all three data sets (Pr = 1.0, BP ⩾ 89).

Although monophyly appeared to be clearly rejected for *Chrysiptera,* the final well-sampled genus falling in Pomacentridae Clade 1, a tree constrained by the monophyly of *Chrysiptera* was not significantly worse than the trees shown in Figs. 2–4 (106mtDNA: $P = 0.1025$, 55RAG1: $P = 0.5969$, and 55combined: $P = 0.3536$). This genus comprises three lineages that we have designated *Chrysiptera* I, II, and III in Figs. 1–3. *Chrysiptera* I includes six species that always group together with high statistical support (Pr = 1.0, BP = 100), and the combined data support a sister group relationship with *Dischistodus melanotus* (Pr = 1.0, BP = 85). However, the other two data sets provide less support for this sister group relationship (106mtDNA: Pr = 0.91, BP = 51 and 55RAG1: Pr = 0.78, *BP* < 50). *Chrysiptera* II includes three species that are strongly associated (Pr ⩾ 0.97, BP ⩾ 80), and fall in a clade including *Neopomacentrus* species but as before, statistical support for this sister group relationship is uneven across data sets (106mtDNA: Pr = 0.47, BP<50, 55RAG1: Pr = 1.0, BP = 66; and 55combined: Pr = 1.0, BP = 85). *Chrysiptera* III includes only *Chrysiptera annulata* in our analyses, and its placement across the different trees presented in Figs. 2–4 is very unstable. Given that we have sampled roughly 30% of the *Chrysiptera* species, our analyses cast doubt that increased species sampling would render the genus monophyletic.

Beyond the associations between Pomacentridae Clade 1 genera noted above, the relationships in this group are not clearly defined, as can be determined from the different phylogenetic placement of lineages in Figs. 2–4. For example, *Pomacentrus* and *Amphiprion* + *Premnas* clades have a different relative placement in each of the three Bayesian trees (Figs. 2–4). More generally an SH test indicated that the topological differences among the trees pictured in Figs. 2–4 were not significant (minimum P value observed = 0.25).

Overall, for the 10 Pomacentridae genera that have been sampled for more than one species, our analysis supports monophyly for seven (*Abudefduf*, *Amblyglyphidodon*, *Dascyllus*, *Neoglyphidodon*, *Neopomacentrus*, *Plectroglyphidodon*, and *Pomacentrus*) and rejects monophyly for three (*Amphiprion*, *Chromis*, and *Chrysiptera*). A tree constraining the monophyly of all Pomacentridae genera was rejected using the SH test (106mtDNA: $P = 0.0004$, 55RAG1: $P = 0.0124$, and 55combined: $P = 0.0001$).

Finally, our analysis of a nuclear gene in comparison to the mitochondrial genes used here and by previous investigators inferring relationships among closely and distantly related Pomacentridae species permits an assessment of the phylogenetic signal provided by the two classes of molecular marker. In particular, there is a striking contrast between the mitochondrial and RAG1 data sets when one compares the level of support for the basal relationships among the Pomacentridae. In the 55RAG1 phylogeny (Fig. 3) the branches relating Pomacentridae Clades 1–4 are all supported with Pr = 1.0 and BP = 100, whereas the mitochondrial tree (Fig. 2) has much lower values. For example the branch uniting Pomacentridae Clades 1 and 2 in the mitochondrial tree is only supported with Pr = 0.61 and BP = 67. The reason for the discrepancy in support values is at least partly explained by the saturation of mtDNA nucleotide substitutions among divergent Pomacentridae lineages. Fig. 1 plots uncorrected pairwise distances of the ncDNA sequences against mtDNA data for the 55 individuals included in the combined data set. At roughly 2% RAG1 sequence divergence, the phylogenetic signal of mtDNA sequences starts to saturate. Saturation is relatively high at genetic distances corresponding to basal relationships within each of Pomacentridae Clades 1–4, and mtDNA saturation is very pronounced for comparisons between species representing the different principal Pomacentridae clades. For RAG1, a graph comparing corrected versus uncorrected genetic distances revealed no indication of

saturation and plots reconstructed separately for each mtDNA genes (ATPase 8, ATPase 6, and cytochrome *b*) indicated saturation above 2% RAG1 sequence divergence in all cases (results not shown).

4. Discussion

The research presented here continues the molecular systematic investigation of Pomacentridae initiated by Tang (2001) and Jang-Liaw et al. (2002) through an increase in the number of taxa analyzed ($N = 103$), inclusion of the RAG1 nuclear gene (1500 bp), and a redirected mitochondrial focus centering on protein coding genes (cytochrome *b* + ATPase 8/6: 1989 bp). We begin our discussion by presenting the limitations of our taxon and nucleotide sampling, in order to permit the reader to better assess the quality of the hypothesis of Pomacentridae relationship that we present in the final section.

4.1. Taxonomic sampling of the Pomacentridae

Our sample of Pomacentridae species and genera is reasonable but nonetheless represents only 103/355 (29%) of the species, 18/29 (62%) of the genera, and three of the four subfamilies (75%) listed for the family by Allen (1991). From a conventional perspective the most critical missing species is *Lepidozygus tapeinosoma*, the monotypic representative of the subfamily Lepidoziginae (Allen, 1975, 1991). Intuitively subfamilies are relatively old lineages, and accordingly one expects that *L. tapeinosoma* would form a long branch relative to other Pomacentridae clades. Owing to the phenomenon of long-branch attraction (Hendy and Penny, 1989), a monotypic subfamily such as the Lepidozyginae needs to be treated cautiously in phylogenetic analyses. Given that we have not sampled the lineage, our concern is not long-branch attraction; rather it is the likelihood that inclusion of *L. tapeinosoma* would break a long internal branch in the phylogenetic hypothesis presented for the Pomacentridae. This is because introduction of a species or a clade that joins a long branch has the potential to significantly affect the support value (confidence estimate) of relationships among taxa subtended by the broken branch (Lecointre et al., 1993; Milinkovitch et al., 1996).

Thus, the absence of *L. tapeinosoma* begs the question of its placement in the tree and potential for disrupting our inference of Pomacentridae relationships. Norman (1957) included *Lepidozygus* in the Chrominae, indicating that this species has the potential to disrupt support for Clade 2 in our treatment of the Pomacentridae. If one informally converts Allen's (1975) provisional key of the Pomacentridae subfamilies to a hypothesis of relationship, the key phenotypic characters would appear to unite the Lepidoziginae with the Pomacentrinae. In this scenario, *Lepidozygus* has the potential to disrupt support for Clades 1, 3 or 4, given that each includes taxa included in Allen's Pomacentrinae, and thus we cannot predict where the species will fall. The almost ubiquitous potential placement of *Lepidozygus* is a concern regarding the overall stability of the Pomacentridae hypothesis presented here. This concern is mitigated by the fact that the species can only fall in one place in the tree and thus its influence on the resulting topology will be restricted. Furthermore, the numbers of nucleotide synapomorphies identifying each of the four principal Pomacentridae clades named in this study are sufficient in number to render unlikely their concerted conversion to homoplasic states following the addition of *Lepidozygus*.

Similar concerns regard the potential phylogenetic placement and disruptive behavior of the 10 other Pomacentridae genera that are not included in our analyses. It is worth noting that none of the missing genera include more than two species: *Amblypomacentrus* (2 spp.), *Altrichthys* (2 spp.) *Azurina* (2 spp.), *Cheiloprion labiatus*, *Hypsypops rubicundus*, *Mecaenichtys immaculatus*, *Nexilosus latifrons*, *Pristotis* (2 spp.), *Similiparma hermani*, and *Teixeirichthys jordani*. Through reference to Tang (2001) and Jang-Liaw et al. (2002) we have insight into the phylogenetic placement of *H. rubicundus*, *M. immaculatus*, and *T. jordani*. *H. rubicundus* would almost certainly fall in our Clade 4, although its sister group relationship is uncertain. In Tang's study *H. rubicundus* appears most closely related to *Parma*, a result consistent with its former inclusion in *Parma* (see Tang, 2001, p. 598). Regarding *Teixeirichthys*, reference to Jang-Liaw et al. (2002) indicates a basal relationship with Clade 1, and thus the species has the potential to destabilize our hypothesis of relationship by breaking the long branch leading to Clade 1. We consider it more likely that the lineage represented by *Teixeirichthys* formed during the early and expansive diversification manifest in the basal polytomy of Clade 1. The very strong statistical support for the monophyly of Clade 1, coupled to the generally unresolved relationships among Clade 1 genera, indicates that the inclusion of *Teixeirichthys* is unlikely to have a significant impact on the general topological attributes of the Pomacentridae tree presented here. The placement of *M. immaculatus* is likely to fall deep in the phylogeny, but comparison of the phylogenetic hypothesis of Jang-Liaw et al. (2002) to ours indicates that any destabilizing influence of *Mecaenichthys* is likely to be restricted to Clade 4.

Obviously, systematic analysis of the seven remaining genera is required to determine their influence on the basal relationships of the Pomacentridae phylogeny presented here. In our view, the potential disruptive influence of these genera depends on the probability that

they are relatively old. Unless there has been extensive extinction of related taxa, or species occupy a fraction of their former ranges, it may be that *Azurina*, *Nexilosus*, *Similiparma*, and *Altrichthys* are relatively young as species of these genera have peripheral and local distributions (*Azurina eupalama*: Galapagos Islands, *Azurina hirundo*: Mexican offshore islands, *N. latifrons*: Galapagos to Chile, *S. hermani*: Cape Verde Islands, *Altrichthys curatus*, and *Altrichthys azurelineatus*: Philippines offshore islands). In this case they are most likely to break long terminal branches and would have little potential to disrupt the hypotheses for Pomacentridae relationships presented here.

Lastly, there are 252 additional species not mentioned above that are missing from our study, but all belong to genera that were included in our analysis of Pomacentridae relationships. Unless some of these species represent surprisingly old lineages the likelihood that they will disrupt the basic Pomacentridae topology presented here is probably small. Forty-five missing species represent *Pomacentrus* and 68 are *Chromis*, i.e., 45% of pomacentrid species missing in this study belong to these two genera. We have analyzed 17 of the 85 *Chromis* species and included representatives from the Mediterranean ($N = 1$), Caribbean ($N = 2$), eastern Pacific ($N = 1$), and the Indo-West Pacific ($N = 13$). The two additional *Chromis* species analyzed by Jang-Liaw et al. (2002), *Chromis fumea* and *Chromis analis*, appear related to the species grouped in our Chromis I sub-clade. Given our complete geographic coverage of the genus and the diversity of *Chromis* lineages included in our tree, we consider it unlikely that the inclusion of additional species will weaken support for Clade 2. However, it is readily apparent that the inclusion of additional *Chromis* species will be necessary to interpret the apparent paraphyly of the genus and to better appreciate its relationship to *Dascyllus*.

We have also analyzed 19 of the 64 *Pomacentrus* species collected throughout the Indo-West Pacific range of the genus (Allen, 1991). We sampled endemic Pomacentrus (e.g., *Pomacentrus australis*, Great Barrier Reef; *Pomacentrus leptus*, Red Sea; and *Pomacentrus milleri*, Western Australia), as well as those with broad distributions (e.g., *Pomacentrus pavo*, *Pomacentrus coelestis*, and *Pomacentrus chrysurus*). Jang-Liaw et al.'s (2002) study clearly associates *Pomacentrus auriventrus* with this group. The very strong statistical support for the monophyly of Clade 1, coupled to the generally unresolved relationships among Clade 1 genera, indicates that the inclusion of additional *Pomacentrus* species is unlikely to have a significant impact on the general topological attributes of the Pomacentridae tree presented here.

We have sampled all but four *Abudefduf* species, and almost certainly the extremes of diversity in this group (see Bermingham et al., submitted), and thus can state with confidence that inclusion of the missing species is very unlikely to cause any reduction in the statistical support for Clade 3. In a similar manner the virtually complete species-level sampling of *Dascyllus*, and good samples of *Amblyglyphidodon*, *Amphiprion*, *Chrysiptera*, *Neoglyphidodon*, *Neopomacentrus*, and *Plectroglyphidodon* permits a first impression that increased species sampling across these genera will probably have only very local topological effects on the Pomacentridae tree. We are then left with *Dischistodus*, *Microspathodon*, *Parma*, *Pomachromis*, and *Stegastes*, for which we have sampled 1/7, 1/4, 1/10, 1/4, and 1/37, respectively, of the species diversity in each genus. Owing to the predominance of these genera in Clade 4, it would appear that increasing the representation of species in these genera has the most potential to impact our hypothesis of Pomacentridae relationships. However, Jang-Liaw et al. (2002) included six *Stegastes* species and although phylogenetic resolution is poor in their analysis, there is no indication that improved sampling of *Stegastes* would have a significant effect on the hypothesis of Pomacentridae relationships presented here.

Nonetheless it should be clearly noted through inspection of the phylogenetic placement of species in well-sampled genera, that support for the four principal Pomacentridae clades is strong and statistically significant even in cases where the clades contain genera that may not be natural groups, such as *Chromis* (Clade 2) and *Chrysiptera* (Clade 1).

4.2. Nucleotide sampling and the phylogenetic utility of the mitochondrial ATP synthase and cytochrome b genes and the nuclear-encoded recombinant activating gene

Our choice of mitochondrial and nuclear genes was based on several considerations. We centered our focus on protein-coding genes because we consider that the mechanisms of nucleotide substitution are best understood for this class of genes. The mitochondrial cytochrome *b* and ATP synthase 6 and 8 genes were analyzed because the first is the most widely used gene in molecular systematics (see Kocher and Stepien, 1997 for examples on fish), and the ATP synthase genes hold particular interest for our lab (e.g., Bermingham et al., 1996; Banford et al., 1999; Hunt et al., 2001; Martin and Bermingham, 2000). Furthermore, the cytochrome *b* and ATP synthase genes sit on opposite sides of the mitochondrial genome, thus decreasing the likelihood that a transposition of a mitochondrial sequence would carry all three genes into the nucleus (Bermingham et al., 1996, see also Lopez et al., 1994). Joint analysis of cytochrome *b* and the ATP synthase genes permits strong inference that a functional mtDNA linkage group is in hand (i.e., guards against analysis of mtDNA pseudogenes), and provides a sufficient nucleotide sample (approximately 2000 bp) to estimate the

genealogical history of the mitochodrion (Bermingham et al., 1996).

The single copy nuclear gene, RAG1, was selected because it is found in all vertebrates and has well studied biological functions (Oettinger et al., 1990; Schatz et al., 1989). Moreover, recent studies have indicated that RAG1 has attributes that make it generally useful for phylogenetic inference (e.g., Groth and Barrowclough, 1999; Quenouille and Bermingham, submitted). For example, Groth and Barrowclough, 1999 found that RAG1 provided a high level of phylogenetic resolution across divergent avian taxa, and was characterized by constant base composition, low levels of nucleotide saturation, and a paucity of indels.

Two general and related, but unsurprising, conclusions regarding RAG1 and the mitochondrial markers can be drawn from our phylogenetic analysis of the Pomacentridae. First, the phylogenetic efficacy of the two classes of molecular markers is manifest at very different levels of relationship within the tree. For example, RAG1 provides virtually no phylogenetically-informative characters among closely related Pomacentridae species, whereas the mitochondrial genes provide ample characters for phylogenetic inference among such species. Thus as can be determined for the taxonomically well-sampled Pomacentridae genera (e.g., *Dascyllus*, *Abudefduf*), relationships among closely related species are well resolved on the mtDNA tree, but not on the RAG1 tree, indicating that the RAG1 substitution rate is slow relative to Pomacentridae speciation rate in these cases. Mitochondrial DNA on the other hand has a reduced number of informative changes per unit length of sequence at tree depths consistent with comparisons among genera or deeper. Our comparison of RAG1 and mtDNA protein-coding genes indicates that the nucleotide sample of nuclear genes would need to be roughly one order of magnitude greater than the mitochondrial sample to provide an equivalent number of phylogenetically informative sites among congeneric Pomacentridae species.

The second conclusion is simply that nucleotide saturation provides the explanation for the loss of mitochondrial phylogenetic signal among more distantly-related Pomacentridae species. Mitochondrial nucleotide saturation becomes apparent in the comparison with RAG1 pictured in Fig. 1 at roughly 10% sequence divergence, and approaches the asymptote at approximately 22%. Thus the increase in mtDNA divergence as a function of RAG1 divergence levels out quickly, such that a 4-fold increase in RAG1 divergence (2–8%) corresponds to a 1.5-fold increase in mitochondrial divergence (17–27%). It is noteworthy that third position substitutions in the mitochondrial protein-coding genes saturate at a tree depth that coincides with genus-level relationships, or even among the most divergent species in a genus. Nucleotide saturation masks variation and increases homoplasy, and in turn leads to an underestimate of bootstrap measures of confidence (Zharkikh and Li, 1992). Comparisons of the bootstrap support for the deep nodes in the Pomacentridae RAG1 (Fig. 3) versus the mtDNA tree (Fig. 2) demonstrate the anticipated pattern, with lower support values deep in the mtDNA tree as compared to RAG1.

Nucleotide saturation leads to analytical difficulties when attempting to resolve phylogenetic relationships between taxa separated by genetic distances that exceed the values at which a majority of nucleotide positions that are free to vary have experienced multiple substitutions. Thus, and particularly with mtDNA, there can be problems distinguishing between a polytomy or near-polytomy that results from a species diversification process that is rapid relative to the nucleotide substitution process (a so-called hard polytomy), and one that owes to nucleotide saturation (soft polytomy). Given that many internal nodes in the Pomacentridae mtDNA tree are between species separated by genetic distances that represent fully saturated mtDNA sequences at 2- and 4-fold degenerate sites, it could be reasoned that any lack of internal resolution in the tree resulted from nucleotide saturation. But this explanation fares less well when considering the RAG1 tree, because RAG1 demonstrates no detectable nucleotide saturation at the genetic distances observed between Pomacentridae species. We consider it more likely that unresolved relationships deep in the RAG1 and combined trees represent hard polytomies, and suggest that the Pomacentridae diversification process has been rapid relative to the RAG1 substitution rate in these regions of the tree. Elsewhere, we have estimated a RAG1 evolutionary rate of 0.6 substitutions per million years of divergence (Quenouille and Bermingham, in preparation), suggesting that roughly 5 million years would be required to accumulate three synapomorphies across 1500 bp of RAG1. Felsenstein (1985) demonstrated that parsimony-based bootstrap values greater than 95% are generally reached when a node is supported by three or more unreversed synapomorphies. Thus, a more temporally constrained inference of the relative branching order of many Pomacentridae genera will require a significantly larger nuclear gene sample of nucleotides than presented here.

4.3. Considerations regarding the analytical treatment of Pomacentridae relationships

Phylogenetic analysis of large data sets requires the examination of a very large number of alternative topologies, which typically require computational time and resources beyond the reach of typical molecular systematic investigations. Although heuristic methods or pairwise distance analyses, such as the Neighbor-Joining method, provide reasonably or very fast

phylogenetic approximations, both approaches have significant shortcomings (Swofford et al., 1996). Heuristic searches represent a very limited estimate of phylogeny, and time requirements for generating bootstrap estimates of confidence for a data set the size of the Pomacentridae dramatically reduces the efficacy of such searches. Generating confidence estimates for a NJ tree, on the other hand, is relatively fast, but reducing character data to pairwise distances thwarts efforts to infer ancestral states. Furthermore, distance-based analyses do not lend themselves to the combination of different classes of data.

Thus for a combination of these reasons we turned to a Bayesian statistical approach as our principal method for the phylogenetic analysis of the Pomacentridae (Huelsenbeck et al., 2001a; Larget and Simon, 1999). Like maximum likelihood, the Bayesian method is a model based approach, and is thus explicit regarding the pattern of nucleotide substitution. Most importantly, the Bayesian approach provides probabilities for hypotheses given the data, rather than the probabilities of the data given a hypothesis, as is the case for ML (Lewis, 2001). ML results are thus more difficult to interpret in comparison to the straightforward assessment of Bayesian support for a phylogenetic hypothesis. The relative novelty of the Bayesian approach has limited its inspection in comparison to alternative phylogenetic methods, but such scrutiny is rapidly expanding (e.g., Buckley et al., 2002; Huelsenbeck et al., 2002; Leache and Reeder, 2002).

Given that Bayesian analyses do not accomplish a complete search of tree space, it is worth noting the approach we took to overcome analytical shortcomings that can potentially limit the application of the method, and in turn our inference of Pomacentridae relationships. In order to insure an adequate approximation of the Pomacentridae phylogeny we ran preliminary analyses to estimate when the likelihood value of each Markov chain reached apparent stationarity, and performed the last run for an additional 2–3 million generations to (1) reduce the probability that we had reached only a local optimum (Leache and Reeder, 2002) and (2) improve the reliability in the confidence interval estimates for all phylogenetic model parameters.

The real advantage of the Bayesian approach lies in the statistical assessment of the hypothesis of relationship presented here for the Pomacentridae. Inspection of the tree presented in Fig. 4, representing the combined mitochondrial and nuclear data, permits the satisfying impression that many clades have high Bayesian support values, but this represents only part of the statistical assessment of the Pomacentridae tree. In addition to focusing on the consensus tree, it is important to evaluate the entire sample of trees, particularly if the sample is rich in distinct topologies. Specifically, one needs to evaluate the sample of distinct topologies, in order of declining frequency, required to reach 95% of all sampled trees; in other words the 0.95 credible set of topologies. In our phylogenetic assessment of the Pomacentridae, the 0.95 credible set of trees for the hypothesis presented in Fig. 4 included 302 distinct hypotheses of relationship, of which the most frequent represented only 6.4% of the sampled trees. Such a statistical assessment indicates that considerable work remains before we are likely to attain a highly confident estimate of Pomacentridae relationships, but the contrast between our low confidence in the overall tree versus the high support values for many of the clades neatly pinpoints where attention needs to be focused.

Our ensuing discussion of the Pomacentridae is based on the combined mitochondrial and RAG1 tree pictured in Fig. 4, and as we noted above statistical support for the entire tree is not significant. Nonetheless, many of the clades have strong statistical support, and thus it is the relationships among some pomacentrid lineages that remain uncertain. Achieving a more confident estimate of Pomacentridae relationships will require a considerably larger nucleotide sample of nuclear genes, and directed inclusion of additional Pomacentridae species that might potentially break long branches leading to internal nodes in the phylogeny.

4.4. The molecular systematics of the Pomacentridae based on mitochondrial and nuclear protein-coding genes

The principal molecular systematic conclusion of our investigation is the recognition of four principal Pomacentridae clades. Each of these clades has strong statistical support, and for reasons discussed above our tentative expectation is that these four clades will sustain additional analytical scrutiny. Only in number do the principal clades diagnosed in our study correspond to the subfamilies recognized by Allen (1991). Our phylogenetic results are broadly consistent with the trees published by Tang (2001) and Jang-Liaw et al. (2002), and thus mitochondrial protein-coding and ribosomal genes are providing a similar estimate of Pomacentridae phylogeny as would be anticipated from a single linkage group, and one that is closely matched in broad outline by the RAG1 tree. Levels of statistical support are generally improved in our phylogenetic analysis in comparison to those of Tang (2001) and Jang-Liaw et al. (2002). The difference probably owes to our use of a larger nucleotide sample, and perhaps to increased certainty regarding the alignment of mitochondrial protein-coding genes in comparison to ribosomal RNA genes.

Our phylogenetic analysis rejects the monophyly of the Chrominae and Pomacentrinae, the two most speciose subfamilies recognized by Allen (1975, 1991). *Acanthochromis polyacanthus* is not a Chrominae as posited by Allen (1991), as it clearly groups with Pomacentridae Clade 1, rather than with *Chromis* and

Dascyllus (Clade 2). *Mecaenichthys*, another Chrominae in the 1991 Allen classification, is also not closely associated with *Chromis* and *Dascyllus* according to the molecular systematic analysis conducted by Tang (2001) and Jang-Liaw et al. (2002). The placement of the monospecific genera of *Acanthochromis* and *Mecaenichthys* notwithstanding, support for a *Chromis* + *Dascyllus* clade is highly significant and the relative phylogenetic separation and basal placement of Clade 2 in the Pomacentridae provides strong support for Allen's (1991) general notion of the Chrominae. Still to be phylogenetically placed is *Azurina*, a genus included in the subfamily Chrominae by Allen (1991) and *Altrichthys* a recently described pomacentrid genus that Allen (1999a) considered most closely allied to *Chromis*. *Azurina* comprised two species with narrow distributions in the eastern Pacific. *Altrichthys* is composed of two species with narrow distribution near the Philippines.

The Pomacentrinae as classified by Allen (1975, 1991) is not a natural group according to our analyses. Nonetheless, Clade 1 includes *Pomacentrus*, the type genus for the subfamily, and appears from a molecular perspective sufficiently well distinguished from other clades to stand as a damselfish subfamily. *Parma*, *Microspathodon*, *Stegastes*, and *Plectroglyphidodon* fall into Clade 4, completely apart from all other Pomacentrinae genera. *Abudefduf*, also included in the Pomacentrinae by Allen (1975, 1991), comprises Clade 3 in our analysis. The molecular systematic assessments provided here, and by Tang (2001) and Jang-Liaw et al. (2002), offer a fresh phylogenetic perspective on the large number of genera that have previously been lumped in the subfamily Pomacentrinae, and indicate some clear directions for future research regarding genus-level relationships in the family.

The third subfamily recognized by Allen (1975, 1991) is Amphiprioninae, uniting *Amphiprion* and *Premnas*. Although our analysis provides strong support for this clade, its placement within Pomacentridae Clade 1 indicates that the *Amphiprion* + *Premnas* clade is probably not sufficiently prominent from an evolutionary perspective to warrant placement in a distinct subfamily. With regard to these genera, our analysis failed to support the monophyly of the genus *Amphiprion*, owing to the well-supported relationship between *P. biaculeatus* and *A. ocellaris*. These two taxa were included in the analyses of Amphiprioninae performed by Elliott et al. (1999), Tang (2001), and Jang-Liaw et al. (2002). They are also identified as sister species in the two MP trees Tang (2001) inferred from a combined analysis of 12S and 16S sequences, and in Elliott et al.'s (1999) MP and NJ analyses of cytochrome b sequences. On the other hand, Elliott et al.'s (1999) MP and NJ analyses of 16S sequences recovered a monophyletic *Amphiprion* and *Premnas* came out as the most ancestral Amphiprioninae. In the MP analyses of 12S and 16S sequences performed by Jang-Liaw et al. (2002), the Amphiprioninae are recovered as a monophyletic clade with a basal polytomy of three lineages, two composed of *Amphiprion* species and the third being *P. biaculeatus*. Such inconsistency probably reflects a mixed effect of nucleotide sampling, taxon sampling and reconstruction methods, but we note that the monophyly of *Amphiprion* would not be in discussion by re-classifying the monotypic genus *Premnas* Cuvier, 1817 within *Amphiprion* Bloch and Schneider, 1801, a decision suggested by the apparent difficulty to find robust phylogenetic signal distinguishing *Premnas* from *Amphiprion*.

Neither our study nor those of Tang (2001) and Jang-Liaw et al. (2002) included *L. tapeinosoma*, the monotypic representative of the subfamily Lepidoziginae (Allen, 1975, 1991), and thus we cannot comment on its phylogenetic placement from a molecular perspective. The subfamilial rank of *Lepidozygus* would predict its placement apart from the four principal Pomacentridae clades identified in our study. However, Fitzpatrick's (1992) cladistic analysis of the Pomacentridae based on 15 morphological characters casts some doubt on the phylogenetic distinctiveness of the Lepidoziginae, as she was unable to diagnose autapomorphies for the lineage.

Although our molecular systematic analysis of the Pomacentridae counters the traditional view (Allen, 1975, 1991) regarding higher-order relationships in the family, Bayesian, NJ, and MP analyses strongly support most genera as natural groups. Considering genera for which we have sampled 20% or more of the species diversity, the monophyly of seven receives significant statistical support: *Abudefduf*, *Amblyglyphidodon*, *Dascyllus*, *Neoglyphidodon*, *Neopomacentrus*, *Plectroglyphidodon*, and *Pomacentrus*. To this list we can add *Amphiprion* if we adopt the simple expedient of re-classifying *Premnas* (see above). The monophyly of *Plectroglyphidodon* is thrown into doubt by the recent study of Jang-Liaw et al. (2002) suggesting the paraphyly of this genus with *Stegastes*. However, the relative lack of resolution in the mtDNA RNA trees in comparison to the mtDNA protein-coding gene tree cautions against over interpretation of the Jang-Liaw et al. (2002) result.

Our analysis provides no statistical support for the monophyly of two genera: *Chromis* and *Chrysiptera*. Although *Chrysiptera* species fall in three different places within Clade 1, a tree constrained by the monophyly of *Chrysiptera* was not significantly worse than the phylogenetic hypothesis presented in Fig. 4. However, the power of this test is compromised by the general lack of distinguishing nucleotide characters at the base of Clade 1, a fact that we relate above to the apparently rapid diversification of Clade 1 relative to nucleotide substitution rate. In addition to questions that our analysis raises regarding the monophyly of *Chrysiptera,* our results also point to hypothetical sister group rela-

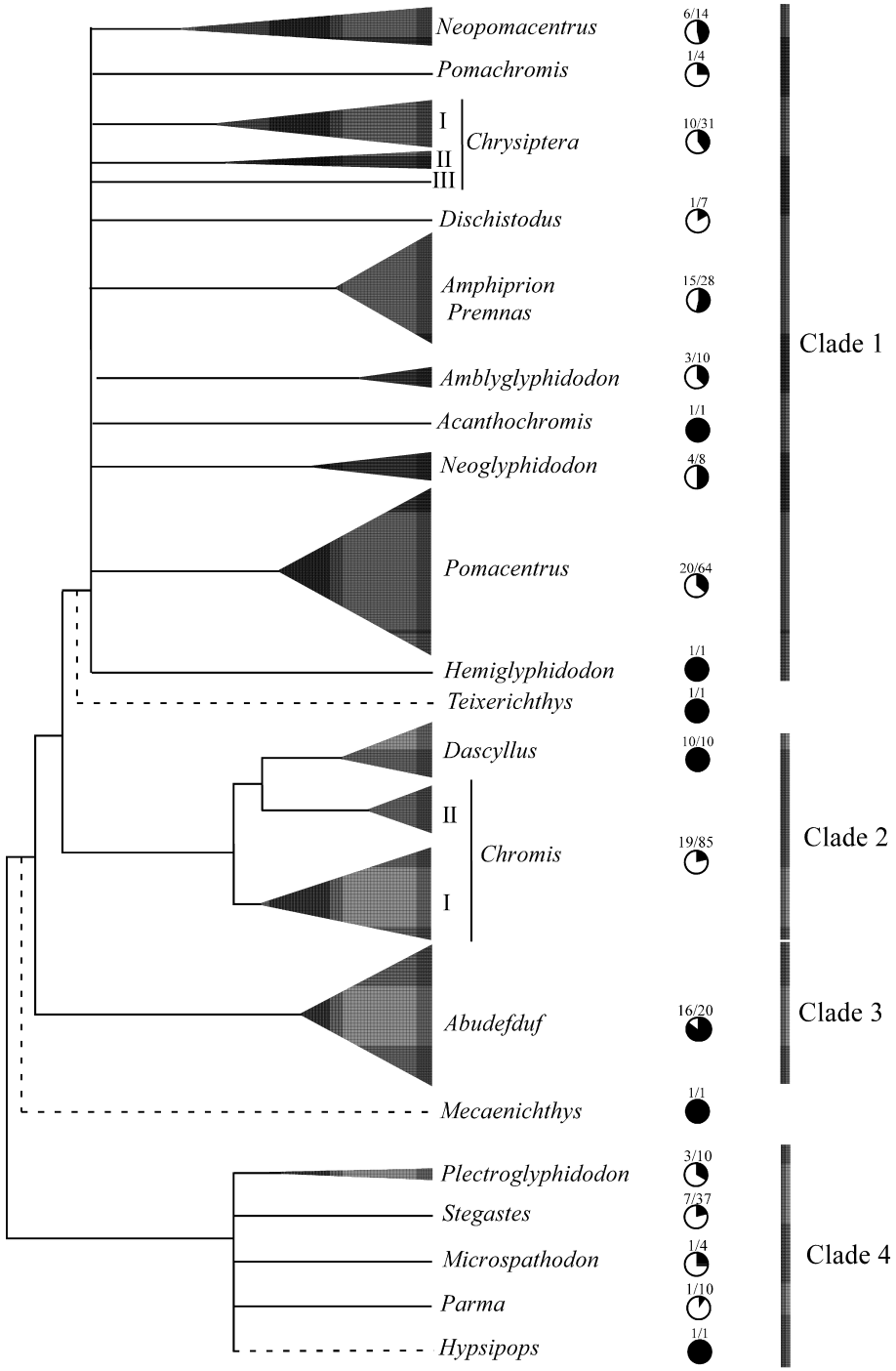

Fig. 5. Phylogenetic hypothesis of relationship among 21 Pomacentridae genera based on a synthesis of the results presented here and by Tang (2001) or Jang-Liaw et al. (2002). Dashed lines identify the putative placement of genera analyzed by Tang (2001) or Jang-Liaw et al. (2002) that were not included in our study. Triangles are proportionally sized to represent the number of congeneric species analyzed in this study (vertical dimension), and to convey the overall level of sequence divergence among congeneric species (horizontal dimension). Pie diagrams and ratios above represent the proportion of congeneric species analyzed in this study in combination with the mtDNA-based studies of Elliott et al. (1999), Tang (2001), Bernardi et al. (2002), McCafferty et al. (2002), and Jang-Liaw et al. (2002) compared to the species counts for the genera.

tionships not anticipated by prevailing taxonomy, e.g., *Chrysiptera* I with *D. melanotus* and *Chrysiptera* II with *Neopomacentrus* (but the statistical support for these sister group relationship is uneven across data sets). Jang-Liaw et al. (2002) documented a well-supported sister species relationship between the only two *Chrysiptera* species that they examined: *Chrysiptera rex* (falling with *Chrysiptera* I in our analyses) and *Chrysiptera brownrigii* (Chrysiptera II). Whether the difference in our results owes to taxonomic sampling, or to

the efficacy of the mtDNA RNA phylogenetic marker in resolving relationships at the base of Pomacentridae Clade 1 would be valuable to determine.

In the case of *Chromis*, a tree constrained by the monophyly of the genus was significantly worse than the phylogenetic hypothesis based on the combined data (as well the mtDNA data alone, but not the RAG1 data). Although our sample of *Chromis* includes only 20% of the described species, the strong statistical support for the clade we designate *Chromis* II and its sister group relationship to *Dascyllus* indicate that increased taxon sampling is unlikely to render *Chromis* monophyletic. Nonetheless, our results indicate that more species and more nucleotides need to be sampled in order to reach a good understanding of the systematic relationships among species currently assigned to *Chromis*, a conclusion not anticipated by some earlier investigators who considered this Pomacentridae genus particularly stable (Allen, 1975; Randall and Swerdloff, 1973). Restoring the predicted stability of *Chromis* requires only the reclassification of *Dascyllus*, a trivial notion but for the fact that taxonomy focuses attention the overall phylogenetic structure and apparent age of clades. For example, Pomacentridae Clades 2 (*Chromis* + *Dascyllus*) and 3 (*Abudefduf*) are strongly supported as natural groups, and although the former contains considerably higher species diversity, the structure and age of the two clades is similar (Quenouille and Bermingham, in prep.). But this apparent fact is not revealed by prevailing taxonomy, and thus the example offered by Clades 2 and 3 indicate that relative changes in the genus-level taxonomy would be required to bring the evolutionary histories of the two clades into better register and focus.

Our molecular phylogenetic analysis of the Pomacentridae indicates a general need for systematic revision of the family and provides useful insight into issues relating to taxon and nucleotide sampling that should help guide future investigations. In Fig. 5, we provide a phylogenetic hypothesis that summarizes our current perspective regarding the molecular systematics of the Pomacentridae, based on the studies of Jang-Liaw et al. (2002); Tang (2001) in combination with our investigation, and representing a total of 120 species and 21 genera. To the extent that the molecular systematic studies of the Pomacentridae are upheld, it follows that characters such as body form, fin ray and scale counts, color, etc., used to infer the current classification scheme for the family are not strong indicators of evolutionary relationship. Furthermore, given that the Pomacentridae represent one of the most intensively studied coral reef fish families, our findings suggest that overall understanding of reef fish systematic relationships is likely to be weak. In our view concerted effort directed at developing character-rich molecular systematic analyses of tropical reef fish would provide the historical data required for sophisticated comparative analyses of the tempo and mode of species production underlying the evolutionary assembly of the most visibly diverse marine community on the face of the earth.

Acknowledgments

B.Q. is particularly grateful to Nimiadina Gomez and Guy Reeves for guiding his first steps through the magic worlds of PCR and DNA sequence analyses. Maribel Gonzalez also provided excellent assistance in the laboratory, and provided high quality DNA extractions from the most recalcitrant tissues. Oris Sanjur's support all along this project was greatly appreciated and with no doubt, deserved the imports of French chocolates. This study could not have been possible without specimen and tissues provided by Kazue Asoh, Tomoko Yoshikawa, Michel Kulbicki, Rene Galzin, Pascal Romans, and Giacomo Bernardi. We thank the people of Kuna Yala for permission to collect in San Blas, and the governments of Panama, France and Australia for permission to collect in their waters. Financial support was provided by the Smithsonian Tropical Research Institute's Molecular Systematics and Evolution program.

References

Allen, G.R., 1975. Damselfishes of the South Seas. T.F.H. Publications, Neptune City, NJ.

Allen, G.R., 1991. Damselfishes of the World. Aquariums Systems, Melle, Germany.

Allen, G.R., 1992. A new species of damselfish (genus *Pomacentrus*) from north-western Australia. Rec. W. A. Mus. 15, 691–695.

Allen, G.R., 1993. Two new species of damselfishes (*Pomacentrus*), with comments on the validity of two additional pomacentrid fishes. Revue Aquariol. 20, 21–26.

Allen, G.R., 1994. Two new species of damselfishes (Pomacentridae) from Indonesian seas. Rev. Aquariol. 21, 86–90.

Allen, G.R., 1999a. *Altrichthys,* a new genus of damselfish (Pomacentridae) from Philippine seas with description of a new species. Revue Aquariol. 26, 23–28.

Allen, G.R., 1999b. Three new species of damselfishes (Pomacentridae) from Indonesia and eastern Papua New Guinea. Revue Aquariol. 25, 99–105.

Allen, G.R., 2002. Description of two new species of damselfishes (Pomacentridae: *Pomacentrus*) from Madagascar. Aqua 6, 45–52.

Allen, G.R., Adrim, M., 1992. A new species of damselfish (*Chrysiptera*: Pomacentridae) from Irian Jaya, Indonesia. Rec. W. A. Mus. 16, 103–108.

Allen, G.R., Adrim, M., 2000. *Amblypomacentrus clarus,* a new species of damselfish (Pomacentridae) from the Banggai Islands, Indonesia. Rec. W. A. Mus. 20, 51–55.

Allen, G.R., Bailey, S., 2002. *Chrysiptera albata,* a new species of damselfish (Pomacentridae) from the Phoenix Islands, Central Pacific Ocean. Aqua 6, 39–43.

Allen, G.R., Rajasuriya, A., 1995. *Chrysiptera kuiteri*, a new species of damselfish (Pomacentridae) from Indonesia and Sri Lanka. Rec. W. A. Mus. 17, 283–286.

Allen, G.R., Randall, J.E., 2002. A review of the *leucogaster* species complex of the Indo-Pacific pomacentrid genus *Amblyglyphidodon*, with descriptions of two new species. Aqua 5, 139–152.

Allen, G.R., Smith, K.N., 1992. A new species of damselfish (Pomacentridae: *Stegastes*) from Ascension Island, Atlantic Ocean. Rec. W. A. Mus. 16, 113–117.

Banford, H.M., Bermingham, E., Colette, B., 1999. Phylogenetic systematics of the *Scomberomorus regalis* (Teleostei: Scombridae) species group: molecules, morphology and biogeography of the Spanish mackerels. Copeia (3), 596–613.

Bellwood, D.R., 1996. The Eocene fishes of Monte Bolca: the earliest coral reef fish assemblage. Coral Reefs 15, 11–18.

Bellwood, D.R., 1997. Reef fish biogeography: habitat associations, fossils and phylogenies. In: Proceedings of the 8th International Coral Reef Symposium, vol. 1, pp. 379–384.

Bermingham, E., Seutin, G., Ricklefs, R.E., 1996. Regional approaches to conservation biology: RFLPs, DNA sequences, and Caribbean birds. In: Smith, T.B., Wayne, R.K. (Eds.), Molecular Genetic Approaches in Conservation. Oxford University Press, New York, pp. 104–124.

Bermingham, E., Banford, H., Martin, A.P., Aswani, V., 1997. Smithsonian Tropical Research Institute neotropical fish collections. In: Malabarba, L. (Ed.), Neotropical Fish Collections, Museu de Ciencias e Tecnologia, PUCRS, Puerto Alegre, Brazil, pp. 37–38.

Bernardi, G., Crane, N.L., 1999. Molecular phylogeny of the humbug damselfishes inferred from mtDNA sequences. J. Fish Biol. 54, 1210–1217.

Bernardi, G., Holbrook, S.J., Schmitt, R.J., Crane, N.L., DeMartini, E., 2002. Species boundaries, populations and color morphs in the coral reef three spot damselfish (*Dascyllus trimaculatus*) species complex. Proc. R. Soc. Lond. B 269, 599–605.

Bleeker, P., 1877. Mémoire sur les Chromides marins ou Pomacentroides de l'Inde archipelagique. Nat. Verh. Holl. Maatsch. Haarlem. 3, 1–166.

Buckley, T.R., Arensburger, P., Simon, C., Chambers, K., 2002. Combined data, Bayesian phylogenetics, and the origin of the New Zealand *Cicada* genera. Syst. Biol. 51, 4–18.

Cuvier, G., Valenciennes, A., 1830. Histoire naturelle des poissons. Tome cinquième. Chez F.G. Levrault, Paris.

Elliott, J.K., Lougheed, S.C., Bateman, B., McPhee, L.K., Boag, P.T., 1999. Molecular phylogenetic evidence for the evolution of specialization in anemonefishes. Proc. R. Soc. Lond. B 266, 677–685.

Emery, A.R., 1973. Comparative ecology and functional osteology of fourteen species of damselfish (Pisces: Pomacentridae) at alligator reef, Florida Keys. Bull. Mar. Sci. 23, 649–770.

Farris, J.S., Kallersjo, M., Kluge, A.G., Bult, C., 1995. Constructing a significance test for incongruence. Syst. Biol. 44, 570–572.

Felsenstein, J., 1985. Confidence limits on phylogenies: an approach using the bootstrap. Evolution 39, 783–791.

Fitzpatrick, S., 1992. Pomacentrid intrafamilial relationships: a cladistic approach. BSC thesis (unpublished).

Gasparini, J.L., Moura, R.L., Sazima, I., 1999. *Stegastes trindadensis* n. sp. (Pisces: Pomacentridae), a new damselfish from Trindade Island, off Brazil. Bol. Mus. Biol. Mello Leitão (N. Ser.) 10, 3–11.

Gluckmann, I., Vandewalle, P., 1998. Morphofunctionnal analysis of the feeding apparatus in four Pomacentridae species: *Dascyllus aruanus*, *Chromis retrofasciata*, *Chrysiptera biocellata* and *C. unimaculata*. Ital. J. Zool. 65, 421–424.

Groth, J.G., Barrowclough, G.F., 1999. Basal divergences in birds and the phylogenetic utility of the nuclear RAG-1 Gene. Mol. Phyl. Evol. 2, 115–123.

Hasegawa, M., Kishino, H., Yano, T., 1985. Dating the human–ape splitting by a molecular clock of mitochondrial DNA. J. Mol. Evol. 22, 160–174.

Hendy, M.D., Penny, D., 1989. A framework for the quantitative study of evolutionary trees. Syst. Zool. 38, 297–309.

Huelsenbeck, J.P., Ronquist, F.R., 2001. Mrbayes: Bayesian inference of phylogeny. Bioinformatics 17, 754.

Huelsenbeck, J.P., Ronquist, F., Nielsen, R., Bollback, J.P., 2001a. Bayesian inference of phylogeny and its impact on evolutionary biology. Science 294, 2310–2314.

Huelsenbeck, J.P., Ronquist, F., Hall, B., 2001b. Mrbayes: A Program for the Bayesian Inference of Phylogeny, Manual.

Huelsenbeck, J.P., Larget, B., Miller, R.E., Ronquist, F., 2002. Potential applications and pitfalls of Bayesian inferences of phylogeny. Syst. Biol. 51, 673–688.

Hunt, J.S., Bermingham, E., Ricklefs, R.E., 2001. The molecular systematics and biogeography of Antillean thrashers, tremblers and mockingbirds (Aves: Mimidae). The Auk 118, 35–55.

Jang-Liaw, N.H., Tang, K.L., Hui, C.-F., Shao, K.-T., 2002. Molecular phylogeny of 48 species of damselfishes (Perciformes: Pomacentridae) using 12S mtDNA sequences. Mol. Phyl. Evol. 25, 445–454.

Kocher, T.D., Stepien, K.L., 1997. Molecular Systematics of Fishes. Academic Press.

Larget, B., Simon, D.L., 1999. Markov chain Monte Carlo algorithms for the Bayesian analysis of phylogenetic trees. Mol. Biol. Evol. 16, 750–759.

Leache, A.D., Reeder, T.W., 2002. Molecular systematics of the eastern fence lizard (*Sceloporus undulatus*): a comparison of parsimony, likelihood, and Bayesian approaches. Syst. Biol. 51, 44–68.

Leão de Moura, R., 1995. A new species of *Chromis* (Perciformes: Pomacentridae) from the southeastern coast of Brazil, with comments on other species of the genus. Rev. Aquariol. 21, 91–96.

Lecointre, G., Philippe, H., Li, H.L.V., Le Guyader, H., 1993. Species sampling has a major impact on phylogenetic inference. Mol. Phyl. Evol. 2, 205–224.

Lessios, H.A., Allen, G.R., Wellington, G.M., Bermingham, E., 1995. Genetic and morphological evidence for an ancient split between the Eastern Pacific damselfishes *Abudefduf declivifrons* and *A. concolor* (Pomacentridae). Copeia (2), 277–288.

Lewis, P.O., 2001. Phylogenetic systematics turns over a new leaf. Trends Ecol. Evol. 16, 30–37.

Linnaeus, C., 1758. Systema Naturae per Regna Tria Naturae Secundum Classes, Ordinus, Genera, Species cum Characteribus, Differentiis Synonymis, Locis. 10th ed., vol. 1. Holmiae Salvii.

Lopez, J.V., Yuki, N., Masuda, R., Modi, W., O'Brien, S.J., 1994. *Numt*, a recent transfer and tandem amplification of mitochondrial DNA to the nuclear genome of the domestic cat. J. Mol. Evol. 39, 174–190.

McCafferty, S., Bermingham, E., Quenouille, B., Planes, S., Hoelzer, G., Asoh, K., 2002. Historical biogeography and molecular systematics of the Indo-Pacific genus *Dascyllus* (Teleostei: Pomacentridae). Mol. Ecol. 11, 1377–1392.

Maddison, D.R., Maddison, W.P., 2000. MacClade 4: Analysis of Phylogeny and Character Evolution, Version 4.0. Sinauer Associates, Sunderland, MA.

Martin, A.P., Bermingham, E., 2000. Regional endemism and cryptic species revealed by molecular and morphological analysis of a widespread species of neotropical catfish. Proc. R. Soc. Lond. B 267, 1135–1141.

Meyer, A., 1993. Evolution of mitochondrial DNA in fishes. In: Hochachka, P.W., Mommsen, T.P. (Eds.), Biology Frontiers, Biochemistry and Molecular Biology of Fishes, vol. 2. Elsevier Science, Amsterdam, pp. 1–38.

Milinkovitch, M.C., LeDuc, R.G., Adachi, F., Farnir, F., Georges, M., Hasegawa, M., 1996. Effects of character weighting and species sampling on phylogeny reconstruction: a case study based on DNA sequence data in Cetaceans. Genetics 144, 1817–1833.

Myers, R.F., 1991. Micronesian Reef Fishes, second ed. Coral Graphics, Guam.

Nelson, J.S., 1994. Fishes of the World, third ed. Wiley, New York.

Norman, J.R., 1957. A draft synopsis of the orders, families and genera of recent fishes and fish like vertebrates. Trustees of the British Museum, London.

Novelli, R., Numan, G.W., Lima, N.R.W., 2000. A new species of the damselfish genus *Stegastes* Jenyns, 1842 (Teleostei: Pomacentridae) from the coast of Brazil. Bol. Mus. Nac., N.S., Zool. 413, 1–12.

Oettinger, M.A., Schatz, D.G., Gorka, C., Baltimore, D., 1990. RAG-1 and RAG-2, adjacent genes that synergistically activate V(D)J recombination. Science 248, 1517–1523.

Palumbi, S.R., 1996. Nucleic acids II: the polymerase chain reaction. In: Hillis, D.M., Moritz, C., Mable, B.K. (Eds.), Molecular Systematics. Sinauer Associates, Sunderland, MA, pp. 205–247.

Posada, D., Crandall, K.A., 1998. ModelTest: testing the model of DNA substitution. Bioinformatics 14, 817–818.

Randall, J.E., 1994. Two new damselfishes waters (Perciformes: Pomacentridae) from Arabien waters. Revue Aquariol. 21, 39–48.

Randall, J.E., 2001. Four new Damselfishes (Percifromes: Pomacentridae) from the Marquesas Isalnds. Copeia (1), 92–107.

Randall, J.E., 2002. Two new damselfishes of the genus *Pomacentrus* from the south-west Pacific. Aqua 5, 167–176.

Randall, J.E., Earle, J.L., 1999. *Abudefduf conformis* and *Plectroglyphidodon sagmarius*, two new damselfishes (Pomacentridae) from the Marquesas Islands. Cybium 23, 333–343.

Randall, J.E., McCosker, J.E., 1992. Two new damselfishes of the genus *Chromis* (Perciformes: Pomacentridae) from the South Pacific. Proc. Calif. Acad. Sci. 47, 329–337.

Randall, J.E., Swerdloff, S.N., 1973. A review of the Damselfish genus *Chromis* from the Hawaiian Islands, with descriptions of three new species. Pac. Sci. 27, 327–349.

Ricklefs, R.E., Schluter, D., 1993. Species Diversity in Ecological Communities. The University of Chicago Press, London.

Robertson, D.R., 1998. Do coral reef fishes have a distinctive taxonomic structure? Coral Reefs 17, 179–186.

Schatz, D.G., Oettinger, M.A., Baltimore, D., 1989. The V(D)J recombination activating gene, RAG-1. Cell 59, 1035–1048.

Shimodeira, H., Hasegawa, M., 1999. Multiple comparison of log-likelihoods with applications to phylogenetic inference. Mol. Biol. Evol. 16, 1114–1116.

Swofford, D.L., 2002. PAUP*, Phylogenetic Analysis Using Parsimony (*and other Methods), Version 4. Sinauer Associates, Sunderland, MA.

Swofford, D.L., Olsen, G.J., Waddell, P.J., Hillis, D.M., 1996. Phylogenetic inference. In: Hillis, D.M., Moritz, C., Mable, B.K. (Eds.), Molecular Systematics. Sinauer Associates, Sunderland, MA, pp. 407–514.

Tamura, K., Nei, M., 1993. Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. Mol. Biol. Evol. 10, 512–526.

Tang, K.L., 2001. Phylogenetic relationships among Damselfishes (Teleostei: Pomacentridae) as determined by mitochondrial DNA data. Copeia (3), 591–601.

Yang, Z., Goldman, N., Friday, A., 1994. Comparison of models for nucleotide substitution used in maximum-likelihood phylogenetic estimation. Mol. Biol. Evol. 11, 316–324.

Zharkikh, A., Li, W.H., 1992. Statistical properties of bootstrap estimation of phylogenetic variability from nucleotide sequences: I. Four taxa with a molecular clock. Mol. Biol. Evol. 9, 1119–1147.